

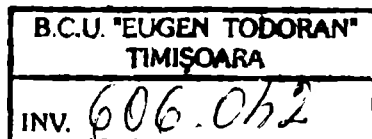
RADU M. SOLCAN

**INTRODUCERE ÎN  
FILOSOFIA MINȚII  
DIN PERSPECTIVA ȘTIINȚEI COGNIȚIEI**

**Editura Universității din București  
2000**

**RADU M. SOLCAN**

**Introducere în  
filosofia minții  
din perspectiva științei cogniției**



Biblioteca Centrală  
Universitară  
Timișoara



02178660

**Editura Universității din București  
2000**

**Referenți științifici: Prof. dr. Adrian Paul ILIESCU**  
**Lector dr. Romulus BRÂNCOVEANU**

**© Editura Universității din București**  
**Șos. Panduri 90-92, București - 76235; Tel./Fax: 410.23.84**

**Descrierea CIP a Bibliotecii Naționale**  
**SOLCAN, RADU M.**

**Introducere în filosofia minții / Radu M. Solcan -**  
**București, Editura Universității din București, 2000**

**268 p.; 23,5 cm.**

**Bibliogr.**

**ISBN 973-575-432-0**

# CUPRINS

FILOSOFIA MINȚII CA FILOSOFIE TEORETICĂ ȘI FILOSOFIA MINȚII	
CA FILOSOFIE APLICATĂ -----	9
Intențiile cărții-----	12
Cum poate fi citită cartea? -----	13
Știința cogniției -----	15
Filosofia minții ca filosofie teoretică -----	17
Filosofia minții ca filosofie aplicată -----	18
Despre unele chestiuni terminologice și nu numai -----	21
PSIHOLOGIE FILOSOFICĂ-----	25
Dualismul cartezian -----	30
Diagrama ideilor -----	31
Dialog pe teme din Descartes! -----	32
Experimentul lui Ryle -----	35
Tipurile și ideea de greșeală categorială-----	36
Fantoma din corp-----	37
Despre minte ... în limbajul obișnuit-----	37
Limbajul comportamentului -----	38
O discuție despre Gilbert Ryle -----	40
Cum descoperiți o temă interesantă?-----	41
De la Descartes la Ryle-----	43
Ludwig Wittgenstein și ideea tipurilor de expresii -----	44
Problema limbajului privat -----	45
După Wittgenstein -----	47
Problema raportului corp-minte -----	48
Funcționalism și complexitate -----	50
Experimentul lui Frank Jackson -----	52



De la fizicalism la <i>qualia</i> .....	53
Un divorț cu acordul ambelor părți.....	55
Și totuși geniul cel rău?.....	57
Filosofie și psihologie .....	61
<b>UNELTELE CELE NOI ALE ÎMPĂRATULUI? .....</b>	<b>63</b>
Filosofie și lingvistică .....	66
Mașini și reguli .....	69
Mașina universală .....	72
De ce Turing? .....	74
Cine calculează?.....	75
Rastere.....	76
Dispozitive simple de prelucrat liste de biți și memorii .....	76
Dispozitive care declanșează alte dispozitive .....	79
Sistemele complexe de prelucrat biți .....	80
Arhitectura computerelor .....	82
Limbaj-mașină cu față umană.....	83
Sisteme de operare și mașini virtuale.....	91
Inteligența artificială .....	93
Algoritmi și structuri de date.....	98
Funcțiile și recursia .....	104
Proceduri, obiecte și modularitate.....	107
Cum să facem față complexității? .....	109
Ce recunosc ușor filosofii printre limbajele de programare? .....	110
Sistemele expert .....	116
Cunoaștere generală și cunoaștere încorporată în cazuri .....	118
Rețele de crezăminte.....	120
Viața artificială.....	123
A fi în lume.....	130
Sistemele artificiale: de la confruntări la rezolvări practice .....	134
Criticii sistemelor artificiale.....	135

Și totuși atracția abordărilor computaționale .....	143
<b>ABORDĂRI COMPUTAȚIONALE ALE MINȚII</b> .....	<b>145</b>
Vederea ca un proces computațional .....	147
Filtrarea critică a ipotezelor cu privire la vedere .....	149
Problemele vederii .....	150
Schița primară: pierdere de date și câștig de structurare .....	151
Modularitatea văzului .....	152
În loc de rezumat: rolul cunoașterii în cadrul vederii .....	154
Structurile de date și memoria .....	155
Experimentele lui Wanner și Marantsos: punerea în evidență a structurilor tip stivă .....	157
Limbajul în perspectivă computațională .....	161
Motorul semantic .....	165
Provocarea ascunsă .....	171
În loc de rezumat: și "mașină" (de calcul) este tot o metaforă .....	176
Modularitatea minții .....	177
Cum de au gândurile efecte în lumea fizică? .....	180
Conexionism contra computaționalism .....	184
Competiția ipotezelor .....	187
<b>NOI ORIZONTURI</b> .....	<b>189</b>
Psihologia cognitivă .....	189
Testul crezămintelor false .....	191
Interpretarea minților .....	194
Incapacitatea de a interpreta mințile .....	197
Deschideri către psihanaliză? .....	199
Trei modele ale minții .....	200
<b>MISTERELE MINȚII</b> .....	<b>203</b>
Câteva repere ale reprezentării cunoașterii .....	204
Două concepții despre minte .....	208
Superveniența .....	209
Tipuri de explicații și experimente imaginare cu ființe zombi .....	214

Conștiință fenomenală și inteligență artificială .....	217
<b>ÎN JUNGLA BIBLIOGRAFIEI .....</b>	<b>221</b>
Bibliografii, ghiduri, dicționare și enciclopedii .....	223
Antologii și introduceri .....	227
Probleme tradiționale în filosofia minții .....	236
Sistemele computaționale și fundamentele lor teoretice .....	241
Mașinăria minții .....	249
Viziunea computațională asupra minții: dezbaterile critice .....	254
Reprezentări mentale .....	257
Nativism, modularitate și limbaj al gândirii .....	259
Interpretarea minților .....	261
Conștiința .....	263

## Filosofia minții ca filosofie teoretică și filosofia minții ca filosofie aplicată

Văd semnele de pe foaia de hârtie. Îmi amintesc că m-am angajat să scriu o carte despre "filosofia minții". Înțeleg textele din antologia pe care o consult. Pentru asta am învățat, între altele, limba engleză. Aș dori să citesc nruvelele polițiste de pe polița din fața mea. Sunt conștient că încerc să vă introduc în ... filosofia minții. S-ar putea toate acestea petrece fără ca "mintea" mea să fie implicată? Nu cumva aceste întrebări sunt însă prea naive, extrem de criticabile din punct de vedere filosofic? Iar tot ce ține de "minte" ar trebui scos din vocabularul pe care-l folosim în mod serios. Sau poate presupoziția că ar exista o minte este falsă?

Ar fi posibil să folosim în loc de "minte" termeni precum "psihic", "spirit" sau "suflet". "Filosofia psihicului" sună însă cam ciudat în limba română. Are un aer de prețiozitate. Altfel, sintagma aceasta este precisă. În plus, subliniază legătura cu *filosofia psihologiei*, studiul filosofic al conceptelor, teoriilor și metodelor din psihologie. Marea problemă pare a fi lipsa intuițiilor care sunt asociate cu termenii curenți în limbă.

"Spirit" este un cuvânt adesea folosit. Eu l-aș utiliza însă în contrast cu "materie". Nu vreau să discut aici meritele sau defectele distincției *materie-spirit*. Doresc doar să păstrez o deosebire între distincțiile filosofice *materie-spirit* și *corp-minte*. În plus, este vorba de a menține și o diferență față de discursul pe teme legate de "spiritual" sau "spiritualitate".

În altă ordine de idei, trebuie amintit că literatura actuală de filosofia minții este în cea mai mare parte scrisă în engleză. A-l reda pe *mind* cu "spirit" este nefiresc. Am călca o regulă euristică simplă: aceea de a nu ne reduce în continuare posibilitățile de mișcare. Când, în engleză, se folosește *spirit*, ar

trebui să punem în română tot “spirit”. Sunt limbi care nu au de ales (cum este franceza, de pildă, unde nu există decât *esprit* sau germana, unde doar *Geist* apare ca o soluție acceptabilă). Italianii preferă pe *mente* și acești cred că este exemplul de urmat. “Suflet” ar fi și mai puțin potrivit, căci el redă pe englezescul *soul*.<sup>1</sup>

Singura obiecție substanțială la adresa termenului “minte”, din câte am auzit, mi se pare cea legată de lipsa dintre sensurile termenului “minte” a celor care trimit la percepții sau trăiri afective. Dacă vom consulta *Dicționarul limbii române*,<sup>2</sup> vom descoperi însă densitatea înțelesurilor lui “minte”: (1) facultatea de a gândi;<sup>3</sup> (2) intenție,<sup>4</sup> atenție,<sup>5</sup> imaginație,<sup>6</sup> memorie<sup>7</sup>; (3) inteligență;<sup>8</sup> (4) înțelepciune;<sup>9</sup> (5) cunoștință<sup>10</sup>.

“Spirit” are și el un spectru larg de utilizări în limba română:<sup>11</sup> este opus materiei<sup>12</sup>, conștiință, cuget, inteligență, fantezie, dar alunecă și spre sensul de *persoană*<sup>13</sup>, alteori de caracteristică sau fond de idei și sentimente ale unei colectivități,<sup>14</sup> ale unei perioade<sup>15</sup>. Ceea ce frapază este sărăcia expresiilor idiomatice specifice *psihologiei populare* (teoriei

<sup>1</sup> Doar redarea tradițională a titlului scrierii lui Aristotel despre psihic este “Despre suflet”.

<sup>2</sup> *Dicționarul limbii române* [prescurtat DLR], tomul VI, fascicula a 8-a, București, Editura Academiei, 1967, pp. 566-570.

<sup>3</sup> Sens vizibil mai ales când ne referim la lipsa de minte (**a nu-și fi în minte**).

<sup>4</sup> **A avea în minte.**

<sup>5</sup> **A fi dus cu mintea** (cu sensul de “a nu fi atent”).

<sup>6</sup> A-i trece (a nu-i trece) cuiva ceva **prin minte**.

<sup>7</sup> **A ține minte.**

<sup>8</sup> **A nu-l ajuge mintea** sau, când este ceva simplu de înțeles, spunem că este **la mintea cocoșului**.

<sup>9</sup> **Mintea cea de pe urmă** sau **cu minte bună**.

<sup>10</sup> **A se învăța minte** (cu sensul de a trage o învățătură).

<sup>11</sup> DLR, tomul X, partea a 4-a, București, Editura Academiei, 1992, pp. 1378-1382.

<sup>12</sup> Poate fi “obiectiv” sau “subiectiv”.

<sup>13</sup> Când spunem despre cineva că este **un spirit elevat**, de pildă.

<sup>14</sup> **Spirit național.**

<sup>15</sup> **Spiritul epocii.**

naive a psihicului prezentă la fiecare om normal, care explică acțiunile, ideile, sentimentelor celorlalți).

“Suflet”<sup>16</sup> este utilizat pentru a ne referi la: viață ca atare,<sup>17</sup> în opoziție cu corpul, sediul sentimentelor (sinonim cu inimă), trăsături de caracter,<sup>18</sup> cuget, conștiință, dar și cu sensul de persoană sau de esență a ceva.

Dacă tragem linie, observăm imediat că “minte” este folosit în contexte în care explicăm comportamentele indivizilor.<sup>19</sup> Dezavantajul lui “spirit” constă tocmai în faptul că nu-l folosim în mod obișnuit pentru a explica ceea ce face cineva. Nu vom zice că “a avut în spirit”, nici că “nu l-a dus spiritul”. “Minte” acoperă foarte bine ceea ce azi se numește în filosofie *minte psihologică*<sup>20</sup>. Evident, “suflet” s-ar potrivi cu ceea ce filosofi numesc *minte fenomenologică* (trăirile calitative). Ralea își dorea chiar o *psihologie a sufletului*. După cum se va vedea în continuare, mai ales din perspectiva științei cogniției, ceea ce putem produce în mod sigur sunt explicații ale *minții* în sens strict. Trăirile sunt în zona misterelor minții. În rezumat, avem nevoie de un termen *uzual* pe care să-l putem apoi asocia treptat cu concepte tot mai rafinate, în contextul unor expresii tehnice. Fără să fie perfect, cel mai bun candidat este cuvântul “minte”.

Mintea nu constituie însă obiect de studiu doar pentru filosofie. Psihologia este știința minții (dacă echivalăm “mintea” cu “psihicul”). Lingvistica este și ea interesată de minte, fie și numai din perspectiva învățării limbajului (de către copii sau adulți). Lor li s-a adăugat, nu cu multe decenii în urmă, inteligența artificială, a cărei țintă o reprezintă scrierea de programe pentru computer, care să se comporte la fel ca și

<sup>16</sup> Vezi DLR, tomul X, partea a 5-a, București, Editura Academiei, pp. 1937-1944.

<sup>17</sup> **A rămâne cu sufletul în oase** (adică a rămâne în viață).

<sup>18</sup> Cineva poate fi **cu suflet** sau **fără suflet**.

<sup>19</sup> Dacă o persoană **și-a adunat mințile** înseamnă că s-a concentrat. Acționăm pentru că **avem ceva în minte**.

<sup>20</sup> Atenție, ce stupid ar suna să spunem “psihic psihologic”! E tot mai limpede acum de ce **nu avem nevoie de un termen tehnic filosofic** aici.

oamenii (să perceapă, să învețe, să înțeleagă limbi naturale ș.a.m.d.). Conceperea ca atare a unor astfel de programe presupune o cercetare și a minții umane, nu doar orientarea către făurirea de minți artificiale.

## Intențiile cărții

Ar fi o banalitate de-a dreptul supărătoare să spui că o introducere în filosofia minții nu-și propune să acopere tot domeniul vizat. Este imposibil să cuprinzi atâtea informații într-un număr redus de pagini. Cu adevărat importantă este perspectiva adoptată. Ea va determina și selecția temelor abordate.

Filosofia are o istorie de aproape două milenii și jumătate. Pentru mulți filosofi două-trei sute de ani nu reprezintă un interval prea mare de timp - în comparație cu această istorie îndelungată. Ei scriu despre memorie sau percepție exact din această perspectivă. Evenimente recente din știință nu par să-i afecteze.

Introducerea de față adoptă o orientare oarecum opusă. Aici, ceea ce am putea denumi cu un termen cu largă acoperire "știința computerelor", joacă rolul unui catalizator.

Mintile celor din jur se crispează când aud de computere. Încă unul care vrea să ne compare cu aceste mașini lipsite de suflet! Trebuie să mărturisesc că această atitudine pe mine unul mă paralizează câteodată. Că ea este reală mi se pare neîndoielnic. Este suficient să te gândești la prioritatea acordată la noi traduceri cărților lui Penrose: critica științei cognitive a venit înainte ca publicul să se poată dumiri despre ce este vorba. Ca o ironie, Penrose cultivă un fizicalism cum rar mai vezi prin filosofie. Dar se pare că a fi un dispozitiv cuantic sună mult mai bine decât provocarea lui Turing: aceea de a produce programe care să aibă reacții aidoma omului. Dacă ai analiza lucrurile bine, ai vedea că, de fapt, cel care nu este mecanicist e mai degrabă Turing.

Dincolo de constrângeri lumești inerente, am scris totuși cartea dintr-un motiv de-a dreptul paradoxal. Unii cărturari subliniază (fără să greșească prea mult) faptul că filosofia tehnică nu are auditoriu. Tehniciștii se citesc unii pe alții (dacă se citesc; de regulă, parcurg părțile care-i interesează direct și ignoră restul). Dar faptul că "tot nu citește nimeni" este un mare avantaj! A scrie devine un exercițiu de descentrare. Ai nevoie de așa ceva pentru a înțelege filosofic despre ce este vorba într-o problemă sau alta.

Scrisul, în acest sens, este un bun substitut pentru conversație. Cele puse pe hârtie devin un obiect de care te poți distanța, pe care-l poți critica, îl poți reface și așa mai departe.

Conversația rămâne, cred, un exercițiu fundamental pentru familiarizarea cu filosofia. De aceea și este recomandată explicit ca metodă de aprofundare filosofică într-una dintre parantezele din carte.

### Cum poate fi citită cartea?

După cum scriam mai sus, cărțile de acest gen au proprietatea de a putea fi parcurse nu doar de la cap la coadă. Recomandarea mea ar fi chiar ca lectura să înceapă cu ultimul capitol.

O privire aruncată prin jungla bibliografică va fi presupusă în continuare, în acest capitol introductiv (dar nu și în capitolele următoare). Bibliografia poate fi luată și ca *dropping names*, cum se spune în engleză. Pomenești pe toți cei pe care i-ai cunoscut, ca să pari mai important. Scopul acestei acțiuni nu este nici pe departe laudabil, dar aici sper că este inocent. Este o ușă deschisă către un univers. Adnotările încearcă să o faciliteze.

Partea documentară a cărții poate fi citită separat. Chiar și fără a parcurge restul materialului sper că este utilă ca platformă de unde să pornești investigații prin sursele de informare de specialitate. Cu puțină răbdare, chiar și aspectele



istorice ale filosofiei minții ar putea fi descoperite cu ajutorul instrumentelor de pe “platformă”.

Puteți citi capitolul al doilea și, din nou, să vă opriți. Puteți să vă opriți, cu alte cuvinte, la problemele tradiționale.

Evident, din perspectiva subsemnatului, capitolul al treilea este miezul cărții (dimpreună cu capitolul al patrulea). Cel dintâi oferă un set de unelte. Cel de-al doilea propune câteva aplicații care folosesc aceste unelte.

Capitolele doi, trei și patru au făcut obiectul a două cursuri ale mele, la Facultatea de Filosofie, dedicate filosofiei minții, pe de o parte, și impactului computerelor asupra filosofiei, pe de altă parte.

Capitolul al cincilea, în schimb, poate fi perceput ca un set de însemnări la cursurile domnului Radu J. Bogdan. De bună seamă, domnul Bogdan nu are nici o vină pentru erorile care ar putea fi sesizate acolo. Rolul capitolului este să pună în evidență semnificația capacității umane *de a citi* alte minți, de a interpreta, cu alte cuvinte, dorințele, intențiile altora, de a ne face, într-un fel sau altul, o idee despre *mintea* lor. Această capacitate este crucială pentru prezicerea cu succes a comportamentelor celorlalți. Sesizarea importanței acestei capacități a sporit enorm respectul nostru pentru *psihologia populară*, care explică și prezice spontan, naiv funcționarea minții. A crescut, probabil, și respectul nostru pentru termenul “*minte*”, în opoziție cu “*spirit*”.

Capitolul dedicat conștiinței acoperă zona în care se găsesc cele mai multe probleme deschise din filosofia minții. Este, altfel spus, locul unde se trage o linie între ceea ce putem spune, cu oarecare siguranță, despre minte și artiile în care predomină problemele nedelegate.

## Știința cogniției

Multe concepte tehnice din știința computerelor au migrat către filosofie, și nu numai către filosofie, prin intermediul a ceea ce se numește știința cogniției.

Știința cogniției este o investigație multidisciplinară în care se întâlnesc eforturile de cercetare a minții din multiple direcții. Când era în construcție, *The MIT Encyclopedia of Cognitive Science* enumera astfel aceste discipline: (1) inteligența artificială și știința computerelor; (2) cultură, evoluție și științe umane; (3) lingvistică; (4) neuroștiință și biologie; (5) filosofia minții; (6) psihologie.<sup>21</sup> În forma finală, lucrurile arată destul de diferit: (1) filosofie; (2) psihologie; (3) neuroștiințe; (4) inteligență computațională; (5) lingvistică și limbaj; (6) cultură, cogniție și evoluție.<sup>22</sup>

Dintre disciplinele de mai sus, primele cinci alcătuiesc nucleul de bază al științei cogniției. Este oare, pur și simplu, interesul pentru *mintea* cel care le-a făcut să se reunească? Nicidecum. În spate se află un anumit mod de a înțelege mintea. Voi puncta aici momentele cele mai frecvent amintite ale genezei acestei abordări.

Din perspectiva psihologiei, crucială a fost descoperirea structurilor interne bogate ale minții. George A. Miller a publicat, în 1956, un articol celebru, în care arăta, pe baza studiilor experimentale, că memoria noastră nu poate menține simultan "în fața ochilor minții" decât șapte itemi (plus sau minus doi). De pildă, un număr de șapte cifre este cam tot ce putem să contemplăm într-un singur moment. Cum am putea ține minte numere mai lungi? Pentru aceasta trebuie să

---

<sup>21</sup> Acestea sunt secțiunile, în 1998, din *MIT Encyclopedia of Cognitive Science*, <http://mitpress.mit.edu/MITECS> [accesată la data de 28.04.1998]. Trebuie precizat că, atât timp cât era în construcție enciclopedia, articolele ei erau accesibile în mod liber pe Internet în versiune completă. După publicare, enciclopedia este doar parțial liber disponibilă pe Internet.

<sup>22</sup> Introducere la *The MIT Encyclopedia of Cognitive Science*, <http://mitpress.mit.edu/MITECS/introduction.html> [accesată pe data de 12.11.1999].

“împachetăm” informația. Putem reține, să zicem, numere de 49 de cifre, dacă le grupăm câte șapte și fiecare grup capătă o semnificație ușor de reținut (numărul de telefon de la facultate, prima parte a adresei numerice a computerului pe Internet și așa mai departe). Morala este că *datele trebuie structurate* pentru a fi memorate.

Exemplul de mai sus este probabil ușor de reținut. Din punct de vedere istoric, contribuția cunoscutului specialist în neurofiziologia creierului, Karl Lashley are prioritate.<sup>23</sup> Lashley a observat că oamenii au adesea scăpări în vorbire care nu pot fi explicate dacă nu postulăm o planificare a activității lingvistice de către subiect. Dacă vom auzi pe cineva spunând “o stradă pe casa Dristor” (în loc de “o casă pe strada Dristor”) ne dăm seama care a fost *structura* planificată și cum au fost așezate greșit elementele în forma pregătită de către vorbitor sau vorbitoare. La fel ar sta lucrurile dacă aş zice “nu există în *mind* cuvânt pentru română”.

Noam Chomsky, în recenzia sa la cartea lui Skinner *Verbal Behavior*, este fascinat de ideea lui Lashley că producțiile noastre verbale nu sunt simple înlănțuiri seriale. Dacă am pus mai sus *mind* în loc de “română” lucrul acesta s-a petrecut pentru că aveam deja o reprezentare a propoziției și m-am grăbit să plasez cu anticipație un element. Chomsky era convins, din cercetările sale lingvistice, că suprafața unui enunț nu ne poate spune care sunt regulile după care este generat. La suprafață, enunțul este o înșiruire de cuvinte, unul după altul. Pentru a le dispune astfel am însă nevoie de *reguli care-mi arată cum procedez*.

Cercetările acestea, după cum se vede, conduc la ideea unor *structuri de date* și a unor *proceduri de prelucrare* a datelor. Or, acestea sunt concepte fundamentale în știința computerelor. Cercul începe să se închidă în jurul unor noi

---

<sup>23</sup> Exemplele menționate în continuare sunt adaptate după cele menționate în Edward Matthei și Thomas Roeper, *Introduction à la psycholinguistique*, Paris, Bordas, 1988, p. 38, după Shattuck-Hufnagel.

unele conceptuale. Astfel ia naștere o nouă disciplină – *știința cogniției*.

### Filosofia minții ca filosofie teoretică

Prin “filosofie teoretică” înțeleg aici, în sens îngust, filosofia ca analiză conceptuală. În acest caz, filosofia clarifică și reconstituie concepte. Rezultatele cele mai bune s-au obținut în acest fel în teoria cunoașterii și filosofia științei.

Paradigmatică pentru filosofia teoretică în sensul de mai sus este relația filosofiei cu fizica. Ipotezele din fizică au *presupoziții* filosofice. Nu este treaba filosofilor să schimbe aceste presupoziții, ci să le explice. Dar, dacă unele conjecturi filosofice sunt preluate de către fizicieni, se instituie un control experimental asupra lor ele își pierd caracterul filosofic.

În filosofia minții, aș distinge două motive teoretice. Primul ține de centrarea pe *metafizica minții*.<sup>24</sup> Filosofia minții, din perspectiva acestui motiv teoretic, își propune să rezolve chestiuni cum sunt raportul dintre corp și minte (este mintea materială sau nu?), problema altor minți (sunt oare toți cei din jurul meu niște roboți bine făcuți?), problema identității personale (dacă îmi transplantați o inimă, rămân aceeași persoană; dar dacă-mi transplantați un creier? Sau combinați o jumătate de creier de la mine cu o jumătate de la altcineva? Sau îmi transplantați doar amintirile altcuiva?) sau problema locului minții în lume.

Un alt motiv este cel al analizei efectelor cauzale ale minții. O posibilă strategie este cea a luării în serios a *simțului comun*. Fehul în care acționăm poate fi explicat, de pildă, prin

---

<sup>24</sup> John Heil, *Philosophy of Mind*, Londra, Routledge, 1998 ilustrează foarte bine acest demers. Deznodământul cărții este legat de identificarea locului minții în cadrul naturii.

combinații de crezăminte și dorințe, așa cum sunt ele formulate în cadrul pe care-l oferă psihologia simțului comun.

Uneori, autorii care cultivă un model de genul celui de mai sus resping explicit orice fel de interes pentru detaliile de implementare.<sup>25</sup> După ei, filosofia minții nu trebuie să se implice în detaliile studiului mașinării minții. Lucrul acesta trebuie lăsat pe seama științei cognitive. Filosofia cognitivei, în sensul acesta, urmărește să clarifice *esența* mentalului. De aceea, în cărțile acestor autori nu vor apărea chestiuni legate de arhitectura sistemelor inteligente sau diagrame ale rețelelor de neuroni.

Ar mai trebui precizat că, în sens larg, filosofia teoretică înglobează și studiul evantaiului de *unelte* ale analizei filosofice. De aici interesul bine înrădăcinat în filosofie pentru logică. Și în cartea de față este cultivat acest interes pentru *unelte*, atâta doar că este lărgită aria lor, prin includerea instrumentelor specifice științei cognitive.

### Filosofia minții ca filosofie aplicată

În continuare, am să încerc să arăt că separarea, în totalitate, a filosofiei minții de implementare (în știința cognitivei) nu este posibilă.

Situația filosofiei în știința cognitivei este cu totul alta decât cea pe care o vedem în raporturile filosofiei cu științele naturii, în primul rând cu fizica.

Zona în care filosofia și fizica se întâlnesc este cea a *experimentelor gândite*. Problema este că, atâta timp cât sunt doar experimente imaginare, ele nu sunt strict vorbind fizică. Filosofia a fost însă puternic influențată de recursul la experimentele imaginare, în această vecinătate extraordinară.

---

<sup>25</sup> Vezi, de exemplu, David Braddon-Mitchell și Frank Jackson, *The Philosophy of Mind and Cognition*, Oxford, Blackwell, 1996, p. X.

Filosofia minții nu este cătuși de puțin o excepție, din punctul de vedere al recursului la metoda experimentelor gândite. Poate mai mult decât în oricare alt domeniu al filosofiei se recurge la experimente imaginare.

Ceea ce este fascinant în știința actuală a cogniției este transformarea, grație inteligenței computaționale, a acestor experimente în experimente cu sisteme artificiale. Și care ar fi atunci deosebirea față de transformarea unor experimente gândite în experimente reale în fizică? Nu se modifică însuși statutul chestiunii analizate, din problemă filosofică într-una empirică? De ce nu ar avea loc procesul de “desprindere de filosofie”, ca și în cazul științelor naturii?

Daniel Dennett arată foarte limpede de ce răspunsul la aceste întrebări este negativ.<sup>26</sup> Construirea unui sistem artificial capabil să execute sarcini cognitive (de exemplu, a unui robot) este un experiment gândit pentru că atunci când datele de ieșire nu sunt cele așteptate este refăcut sistemul. Într-un experiment real nu putem decât să “secționăm” realitatea. Nu putem să mergem în altă “realitate”, cu speranța că aceasta va asculta de legile fizice pe care noi vrem să le examinăm. Într-un experiment gândit, condițiile în care se desfășoară acesta sunt stabilite în funcție de ceea ce vrem să arătăm.

De ce ar fi însă vorba despre filosofie? Dennett explică rațiunile pentru care este vorba despre filosofie. Sistemele artificiale izolează deliberat cazuri. În acest fel pot fi examinate condițiile minime în care sistemul are o funcție cognitivă sau alta.

Întrebările puse prin aceste experimente sunt foarte *abstracte*. Ajunși în acest, punct putem explica și de ce se întinde cercul disciplinelor care fac parte din știința cogniției și peste filosofie.

---

<sup>26</sup> Daniel Dennett, “Cog as a Thought Experiment”, <http://ase.tufts.edu/cogstud/papers/cogthogt.htm> [varianta finală pentru *Robotics and Autonomous Systems* – 8 decembrie 1995].

Până acum am explicat cum au ajuns atât de importante structurile de date și procedurile de prelucrare a datelor pentru înțelegerea minții și că ele pot fi simulate pe calculator. Pentru a face acest lucru este nevoie însă de programe adecvate. Chiar dacă lucrurile vor deveni mai clare când vom discuta modul în care sunt scrise programele, ideea de principiu este că programul este conceput ca o *abstracție* și abia apoi este construit treptat ca program efectiv. Or, investigarea de abstracții este tocmai terenul filosofiei. Aceasta nici nu are unde să se “mute”, ca să spunem așa. Poate să eșueze, să dispară, dar nu poate evita testul aplicațiilor în acest caz.

Nu este de mirare că sabia testelor atârnată deasupra capetelor creează frustrare. În parte, criticile la adresa abordării computaționale sunt încercări de a arăta că filosofia rămâne o activitate strict teoretică. Dar obiecțiile nici nu își ating ținta, pentru că au drept presuposiție tacită ideea că ar fi vorba despre teste empirice și despre eliminarea filosofiei din joc. Nu este așa. Este vorba despre experimente *gândite*, dar nu imaginare. Criticul este într-o situație incomodă: ar trebui să propună un contraexperiment gândit sau să se retragă din jocul filosofic. Tocmai de aceea cele mai reușite demersuri critice, după cum vom vedea mai jos, implică recursul la experimente gândite.

Adevărata parte dramatică este cu totul alta. Jocul s-ar putea să ne depășească puterile. Faptul că am dezlegat misterele mișcării planetelor sau pe cele ale fizicii (până la un punct) sau pe cele ale biologiei nu înseamnă că suntem în stare să facem același lucru și în cazul minții.<sup>27</sup>

<sup>27</sup> Colin McGinn, într-un context diferit, crede că s-ar putea chiar să existe o demonstrație în acest sens (vezi Colin McGinn, *The Problem of Consciousness*, Oxford, Blackwell, 1991, pp. 1-22. Vezi și Colin McGinn, *Problems in Philosophy*, Oxford, Blackwell, 1993, capitolul 2 (Consciousness). McGinn crede că soluția problemei conștiinței ar fi una *naturală*, dar aflată în afara puterilor noastre. De aceea, problema ne apare drept una *filosofică* (cf. p. 44). Poziția adoptată de noi aici este diferită: problema ține de *filosofie*, dar de *filosofia aplicată*. Desigur, s-ar putea ca soluția să nu existe. Derutant este, într-adevăr, faptul că toate ipotezele posibile par să fi fost încercate, dar nici una nu rezistă.

## Despre unele chestiuni terminologice și nu numai

Din pricina interesului pentru *distincțiile conceptuale*, în filosofie există o mare preocupare pentru alegerea termenilor (alegere care reflectă distincții). Ceea ce vreau să sugerez în continuare este o strategie a focalizării: alegerea termenilor este importantă, dar nu are rost să ne concentrăm simultan asupra tuturor termenilor. În primul rând, nu putem să analizăm toate distincțiile odată. Ca să nu rămânem blocați în chestiuni secundare, strategia focalizării constă pur și simplu în folosirea cu strictețe a termenilor implicați în problema analizată și tratarea *generoasă* a celorlalți.

Care ar fi un exemplu de tratare *generoasă*? Fie cazul perechii de termeni “mental-mintal”. Sunt filosofi (Donald Davidson, de pildă) care admit existența *proprietăților mentale*, dar nu și existența unei *minți* (strict vorbind, mințile sunt minți materiale, pentru unii dintre ei). Dacă insist să folosesc pe *mintal*, atunci trebuie să fiu cel puțin pregătit să explic presupuziția admisă astfel (cea a existenței unei minți). Prefer să fiu prudent și, de câte ori nu cred că este vreo miză specială în joc, să folosesc *mental*, care poate fi înțeles și fără a presupune existența unei minți.

Cazul perechii de termeni “limbaj-limbă” ne pune în fața altei probleme. De această dată, strategia este să acceptăm mai multe presupuziții decât ar fi cazul. Este drept, nu am auzit decât despre “limbaj de calculator”; dar unii țin la sintagma “limbă naturală”, ca la un fel de talisman. Practic, nu vom contrazice aceste obiceiuri, deși opinia noastră ar fi că distincțiile conceptuale ar trebui să fie un răspuns la o problemă. Nu are rost să ținem la ele dacă nu putem pune în evidență dificultatea la care ajungem dacă nu le respectăm. Nu văd la ce încurcătură ajungem dacă spunem “limbaj natural”. Deja calificativul “natural” este, ca să spun așa, o pavăză.

Un caz mai delicat este cel al *computațiilor*. În recente sale reflecții critice, Jerry Fodor dă o formulare concisă versiunii



clasice a definiției unei computații văzută ca relație: (1) între simboluri; (2) cauzală; (3) cu conservarea conținutului semantic.<sup>28</sup> Nu vom încerca să analizăm aici reflecțiile critice ale lui Jerry Fodor. Dificultățile încep însă chiar de la “simboluri”. Ce sunt? Reprezintă ceva? Trebuie analizate cu ajutorul instrumentelor logice? Precizia în exces creează, paradoxal, o mlaștină în care analiza se scufundă fără rost. Ceea ce noi vom presupune în continuare este că sistemul are la intrare capacitatea de a distinge între “da” și “nu”. Asta înseamnă că, pe o listă de proprietăți, se bifează sau nu proprietatea, după cum este cazul. Ceea ce se reține este însă doar lista de “da” și “nu”.

Mai departe, caracterul cauzal al relației ridică și mai multe probleme filosofice. Este oare nevoie de o întreagă teorie a cauzalității? Nici nu dispunem de așa ceva. Tot ce vom presupune este că e vorba despre o cauzalitate care respectă *principiul antropic*: nu face, cu alte cuvinte, din start imposibilă existența minților.

Cât despre “conținut semantic” lucrurile stau, probabil, și mai rău. Ne-ar trebui, de la bun început, o teorie a înțelesului. Or, ea este tocmai în miezul disputelor.

Nici vorbă nu poate fi de un concept strict. Tactica este să întindem o plasă mare în care să prindem și computațiile, chiar dacă nu avem exact idee pe unde. Plasa aceasta este de genul următor: avem de a face cu sisteme în care se desfășoară procese. Aceste procese presupun intrări, prelucrare și ieșiri. Dacă aplicăm principiul generozității, vom vorbi pur și simplu despre *proces*. Pe măsură ce strângem plasa, vorbim despre *proces computaționale*. Desigur, că nu am prins încă într-o doară ce doream, dar apelul la generozitate funcționează în continuare.

Adevărul este că partea interesantă a întregii acțiuni nici nu este țelul final, ci modul în care strângem plasa. Filosofii preocupați de știința cognției au venit cu uneltele de acasă; cu alte cuvinte, cu instrumentele logicii. Intenția centrală a cărții noastre este de a arăta că aceste unelte sunt necesare, dar nu suficiente.

<sup>28</sup> Jerry Fodor, *Concepts*, Oxford, Clarendon Press, 1998, p. 11.

Eu unul m-aş fi aşteptat acum un deceniu, un deceniu şi jumătate ca filosofii să înveţe cu nesaţ programare, aşa cum au învăţat în prima parte a secolului logică. Nu s-a întâmplat aşa, deşi exemple lucide au existat (cred că Paul Thagard *Computational Philosophy of Science* este un astfel de exemplu). Marea problemă este trecerea la filosofia aplicată, în sensul de mai sus, şi obstacolele care stau în calea ei.

Sugestia noastră este că aceste reticente au drept cauză presupuziţiile construcţiei de programe: ceea ce am putea denumi o *perspectivă inginerască*. Principii metodologice ale acestei perspective, precum simplitatea, abstracţia şi modularitatea, joacă un rol semnificativ în cadrul abordării computaţionale. Ele sunt însă prea "moi" probabil, în comparaţie cu principiile "tari" ale logici. Terenul însuşi creat de ele pare prea moale pentru fundaţia unui edificiu filosofic. Se prea poate să fie aşa. Dar este un spaţiu pentru experimente gândite, inclusiv cele cu sisteme artificiale. Modul de a face filosofie se schimbă.

Cartea de faţă este o introducere, o introducere în primul rând într-un atelier filosofic în care unele sunt altele decât cele de acum o jumătate de secol. Pentru filosofie intervalul acesta este evident mic. De aceea şi efectele schimbării se văd mai greu.

Cu alte cuvinte, cartea nu încearcă să răspundă la întrebări de genul "ce este...?", ci la cele de tipul "cum fac...?". Nu există în continuare o viziune elaborată asupra minţii sau a fundamentelor conceptuale ale ştiinţei cognitive.

Dacă aplicăm însă ideea generozităţii, putem răspunde scurt la întrebări de genul "ce este mintea?". Răspunsul ar fi "un proces" (cu intrări şi ieşiri), un proces de genul celui descris chiar la începutul acestui capitol introductiv. Acest proces are o proprietate ciudată: dacă ar vrea cineva să-l pună sub control total, atunci efortul acesta va face să existe chiar procesul respectiv. Acestea fiind zise, putem trece la lecturi din Descartes, filosoful care-şi pune pentru prima oară în acest fel problema minţii.

## **Psihologie filosofică**

Nu este deloc greu de înțeles de ce mintea este un subiect tentant pentru filosofie. Mîntea este, ca să spunem așa, la îndemîină. Speculația filosofică se poate exercita pe un teren direct accesibil. Dacă privesc, pot să-mi dau seama care este conținutul vederii mele: foaia de hîrtie, parcul din depărtare, ecranul calculatorului. Dacă vreau să-mi amintesc ce am făcut ieri, pot să fac asta pe îndelete și să urmăresc cum se leagă gândurile mele, cum diverse amănunte devin chei ale porților reamintirii. Dacă vreau să decid ce voi face mâine, pot examina felul în care cercetez alternativele, cântăresc avantajele legate de facerea unui lucru sau a altuia. Cu atît mai pasionant este să urmăresc depănarea unor gânduri sau a modului în care învăț o limbă naturală ori un limbaj de computer. Toate acestea sunt motive suficiente pentru a face din minte un teren pe care speculația filosofică să se desfășoare nestingherită.

Cînd citesc opere literare, pot lesne retrăi pasiunile personajelor, pot desluși motivele care le fac să acționeze. În sensul acesta, nu doar propriile mele gânduri, trăiri sau amintiri îmi sunt accesibile, ci și multe alte minți, reale sau fictive.

Evident, în viața cotidiană, ne întrebăm curent dacă altora le este frică, se bucură sau dacă au priceput ceva, dacă și-au amintit și așa mai departe. Cu alte cuvinte, transformând într-o metaforă contextul lecturii, invocat mai sus, putem spune că suntem în stare să "citim mintea altora". Această "lectură" o vom adăuga astfel la setul de fenomene care ne provoacă la reflecție.

Meditațiile invocate mai sus au însă un aer pregnant "literar". Este vorba de genul de trăiri sau reflecții pe care ne așteptăm să ni le producă o operă literară. Scriitorii, de altfel, fac adesea operă psihologică. Un autor precum Marcel Proust este un maestru în arta de a ne face să trăim sau să retrăim scenele de viață din perspectiva conștiinței. Gesturi banale, ca acela al

înmuierei unei madlene în ceai, capătă cu totul alte dimensiuni din perspectiva fluxului amintirilor autorului. Iar toată întâmplarea narată de Proust este semnificativă pentru propria noastră viață psihică. Tot așa declanșăm și noi procesul de iluminare a ungherelor memoriei: detalii aparent neînsemnate ne sunt de mare ajutor; ne agățăm de ele precum alpiștiștii de colțurile de stâncă și redescoperim ceea ce era acoperit de uitare. Literatura ne ajută efectiv să ne înțelegem mai bine mintea și psihologia poate fi practică cu mijloace literare.

Ce ar deosebi însă psihologia filosofică de cea literară? La urma urmei, granița pare chiar destul de greu de trasat, pentru că mijloacele folosite și terenul pe care ne mișcăm nu sunt tranșant delimitate.

Așa și este. Delimitările nu sunt categorice. Vom încerca să precizăm însă ce anume constituie, din punctul nostru de vedere, specificul filosofiei minții.

Dacă am menține tonul discursului nostru în nota termenilor vagi, atunci ar trebui, în primul rând, să recunoaștem existența unui interes pentru minte încă în opera clasicilor filosofiei antice, Platon și Aristotel. Iar interesul pentru spirit, în opoziție cu materia, este propriu nu doar filosofiei, ci și teologiei.

Filosofia minții, într-un sens mai precis, trebuie văzută însă din perspectiva distincției minte-corp și problemele ei trebuie separate de cele de genul raportului dintre spirit și materie. Este vorba, la urma urmei, despre corpul omului, al animalelor și, mai nou, al roboților. Mintea este văzută în contrast cu acest corp: poate să fie altceva, poate să fie identică într-un fel sau altul cu corpul și așa mai departe.

Nici această restrângere de perspectivă nu este suficientă însă dacă ne gândim că distincția minte-corp este cumva fixată deja în limba pe care o vorbim. Folosim două cuvinte diferite și părem a le asocia unor concepte diferite. De asemenea, familiile de concepte pe care le-am putea delimita sau asemănările pe care le-am putea degaja sunt diferite.

În nici un caz filosofia minții nu începe o dată cu folosirea în limbile naturale a unor termeni distincți pentru corp și minte. Filosofia minții, în ciuda rădăcinilor pe care le-am putea găsi adânc înfipite în istoria filosofiei, își are originea, în forma ei actuală, în opera lui Descartes.

## Experimentul lui Descartes

La Descartes găsim nu doar distincții conceptuale fundamentale pentru filosofia minții sau o teorie despre minte, ci mai ales, cred eu, o anume artă de a mânui o tehnică extrem de utilă în filosofie: experimentul gândit.

Să deschidem textul *Meditațiilor metafizice*<sup>29</sup>, publicate de către Descartes în latină, mai întâi în 1641 la Paris și apoi, în 1642, la Amsterdam. Titlul complet al ediției de la Amsterdam precizează că Descartes vrea să arate că există o distincție între minte și corp: *Meditationum de prima philosophia in quibus Dei existentia et animae a corpore distinctio demonstrantur*. Titlul ediției pariziene, mai prudent, vorbește despre *animae immortalitatis* (nemurirea sufletului). Să ne îndreptăm atenția către alineatul al doisprezecelea. Iată ce scrie Descartes:

*Supponam igitur non optimum Deum, fontem veritatis, sed genium aliquem malignum, eundemque summe potentem et callidum, omnem suam industriam in eo posuisse, ut me falleret;*

Descartes ne îndeamnă să facem presupunerea (fantastică poate) că suntem în puterea nu a unui Dumnezeu foarte bun, izvor al adevărului, ci a unui geniu rău, deopotrivă

<sup>29</sup> Eu am folosit René Descartes, *Méditations métaphysiques*, ediție bilingvă (text latinesc și franțuzesc) realizată de Florence Khodoss, Paris, PUF, 1988 (prima ediție 1956). Trimiterile la diviziunile textului *Meditațiilor* și nu la pagini sper să ușureze însă folosirea de către cititoare sau cititor a oricărui alt text.

puternic și viclean, care-și folosește toată abilitatea sa pentru a ne înșela.

Desigur, Descartes utilizează persoana întâi singular; el este cel înșelat de geniul rău. Generalizarea, prin folosirea în discurs a pluralului nu face însă decât să sublinieze semnificația experimentului imaginar. Fără această generalizare, experimentul filosofic respectiv ar fi lipsit de forță. Geniul rău este un mod de a pune la încercare toate intuițiile noastre naive despre minte.

Care sunt aceste intuiții naive? Ele sunt chiar acele idei precizate atunci când am început să vorbim despre psihologia filosofică. Cred efectiv, de pildă, că am o foaie de hârtie în față, pe care este un fragment din prima meditație metafizică a lui Descartes. Serviciul pe care mi-l face presupunerea existenței unui geniul rău în acțiune este acela de a mă forța să examinez critic astfel de intuiții naive. Filosofia se distinge, de altfel, de alte preocupări ale minții umane prin această forțare a examenului critic.

Descartes precizează ce anume este supus examenului critic, ce anume ar putea să nu fie altceva decât iluzii create de către geniul cel rău:

*putabo caelum, aërem, terram, colores, figuras, sonos,  
cunctaque externa nihil aliud esse quam ludificationes  
somniaorum, quibus insidias credulitati meae tetendit.*

Parafrazându-l pe Descartes, putem spune că geniul rău mă face doar să cred că există clădirile din jur sau că motanul meu are nasul negru și miaună. Toate acestea sunt iluzii create cu dibăcie de către geniul cel rău.

Să zicem însă că nu avem un corp, ne spune Descartes, și că ideea că avem mâini, ochi și așa mai departe este sădită în noi de către geniul cel rău:

*Considerabo meipsum tanquam manus non habentem,  
non oculos, non carnem, non sanguinem, non aliquem  
sensum, sed haec omnia me habere falso opinantem.*

Aici experimentul filosofic propus de Descartes ajunge la un moment-cheie. De ce? Pentru că, în fond, orice astfel de experiment forțează introducerea în argumentarea noastră a ceva: de pildă, a unei distincții. Experimentul are o funcție critică. Ne ajută să separăm aspecte pe care altfel nu ne este lesne să le gândim separat.

Aici este vorba despre o distincție între *conținuturi mentale* (conținutul minții noastre) și mintea noastră ca atare. Logica, nu faptele ca într-un experiment empiric, este cea care ne forțează decizia. Descartes, conform condițiilor stipulate la începutul experimentului gândit, se hotărăște să accepte provocarea geniului cel rău:

*Manebo obstinate in hac meditatione defixus, atque ita, siquidem non in potestate mea sit aliquid veri cognoscere...*

Cu alte cuvinte, acceptăm, vorbind la modul general, ideea că geniul cel rău a ajuns să-și atingă scopul. Geniul a reușit ca fiecare conținut mental să fie stabilit de el (cel puțin atunci când este vorba despre conținuturi mentale referitoare la obiecte externe). Nu poate însă să ne păcălească și în sensul că ne sădește ideea că avem o minte *a noastră*. Dacă nici mintea nu ar fi a noastră, ci doar o iluzie, atunci activitatea geniului ar fi lipsită de sens. Pe cine ar înșela? Se amăgește pe sine? Ce rost ar avea acest lucru?

Ideea că avem o minte a noastră (pe care o umple de iluzii geniul în experimentul imaginar) poate fi extrasă din ideea lui Descartes că

*... at certe quod in me este ne falsis assentiar, nec mihi quidquam iste deceptor, quantumvis potens, quantumvis callidus, possit imponere ...*

Oricât de puternic ar fi geniul rău, oricât de viclean, tot putem să nu-i acceptăm iluziile. Ca orice minte și mintea supusă experimentului imaginat de Descartes poate respinge. Poate respinge fie adevărul, fie falsul. În situația aceasta, evident, este în stare să nu accepte falsurile. Este acesta chiar semnul că este

o minte independentă de mașinațiile geniului cel rău. Iar Descartes și încheie alineatul cu cuvintele

... *obfirmata mente cavebo*.

*Mens* (mintea) este ceea ce capătă, ca urmare a experimentului filosofic cartezian, un contur precis. Conținuturile mentale pot fi nesigure. Mintea ca atare are însă sigur o existență.

### Dualismul cartezian

Ce fel de existență are însă mintea? Este unul dintre obiectele din lumea fizică? Sau unul dintre procesele din lumea fizică? În ce raport se află cu corpul ?

Descartes a formulat o teorie cu privire la raportul dintre minte și corp. După el, mintea este distinctă de corp. Ea este caracterizată de atribute precum gândirea.<sup>30</sup> Asemenea atribute nu pot fi detașate de mintea noastră (îi aparțin, în sensul în care avem o minte care este a noastră). În cuvintele lui Descartes:

*Sum igitur praecise tantum res cogitans, id est, mens, sive animus, sive ratio ...*

În sensul acesta, noi suntem minți. Mintea, crede Descartes, din motive care au inspirat rândurile scrise în deschidere despre psihologia filosofică, este mai ușor de cunoscut.

Dar corpul? Corpurile fizice în general, socotește Descartes, au atributul întinderii. În termeni ceva mai detașați de textul lui Descartes, lumea corpurilor fizice este o lume de cauze mecanice. Dacă ea există, mintea nu ar putea fi un obiect sau un proces din această lume fizică, deoarece aceasta ar însemna să explicăm prin cauze fizice procesele minții. De aici nevoia de a admite dualitatea corp-minte.

---

<sup>30</sup> Vezi *Meditația a doua*, alineatul 7.



## Diagrama ideilor

Filosofia este prin excelență un joc cu abstracția. Anticipând, putem spune că abstracțiile joacă un rol cheie în construirea de programe și, implicit, în sistemele folosite în experimentele gândite ale filosofiei minții ca filosofie aplicată. Obişnuința de a utiliza abstracții ne este însă bine-cunoscută din operații intelectuale mai modeste, cum ar fi rezumatele, listele cu obiective ale acțiunilor noastre. În toate aceste cazuri, mascăm detaliile (sau chiar unele aspecte ale acțiunii – mijloacele ei, de pildă, sau scopurile ei etc.) pentru a surprinde *liniile de forță*.

Prin urmare, în parantezele pe care le vom face, interesul principal nu stă în decelarea dinamicii de detaliu a textelor sau ideilor, ci în sesizarea liniilor de forță ale câmpului problemelor puse de un text sau o temă de cercetare. Ambiția este de a decela dinamica acestui câmp independent de litera textelor.

Paragrafele tipărite cu litere de genul celor din această porțiune de text descriu tehnici care permit explorarea câmpului problemelor în sensul de mai sus. Aceste paragrafe pot fi omise de către cei care doar vor să vadă despre ce este vorba în filosofia minții sau pot să fie citite selectiv. De exemplu, paragrafele de genul celui de față oferă implicit rezumate ale discuției și pot fi utile oricui. Paragraful care urmează este însă indispensabil doar celor care vor să aprofundeze filosofia minții.

Diagramele sunt un mijloc simplu de a schematiza dinamica unui câmp de probleme. Sub forma unor grafice, ele sunt mai sugestive, dar periculoase în cazul structurilor mai complexe. Capcanele pe care le întind diagramele ne interesează însă mai puțin pe moment. Ceea ce vrem să le propunem cititoarelor și/sau cititorilor este un mijloc de a trasa drumul parcurs între două puncte date și o orientare la răscrucea drumurilor care se deschid în continuare.

Dacă vor, cititorii sau cititoarele pot transforma diagrama construită mai jos utilizând forme geometrice și săgeți. Așa vor descoperi poate mai lesne ce căi poate urma discuția în continuare.

Diagrama construită de către noi are o idee simplă. Ea se compune din blocuri delimitate prin paranteze. Acoladele sunt folosite pentru a marca ceea ce putem denumi metode (tehnicile intelectuale, uneltele folosite de către filosofi pentru a obține rezultate). Parantezele drepte sunt folosite pentru a încadra repere ale pozițiilor filosofice.

Fiecare bloc este precedat de un titlu scris cu litere mai îngroșate, pentru a fi ușor de observat. Trecherile de la un bloc la altul, dacă este cazul, sunt precizate în limbaj natural, ca și conținutul blocurilor.

Diagrama ideilor, până în momentul de față, arată astfel:

- **Experimentul cu geniul cel rău**

*{Un geniu rău încearcă să ne insufle tot felul de crezăminte false. Experimentul forțează distincția între conținuturi mentale și minte.}*

Experimentul conduce la ideea că

- **Mintea**

*[Mintea există ca o entitate aparte.]*

De asemenea, după Descartes, trebuie acceptată

- **Distincția dintre minte și corp**

*[Mintea are o natură diferită de cea a corpului.]*

Oricare dintre cele trei blocuri ale diagramei de mai sus poate avea fisurile sale filosofice. De aceea ne și așteptăm că un filosof sau altul au încercat să le exploateze.

## **Dialog pe teme din Descartes!**

Discutarea legăturii dintre filosofie și conversație este dincolo de obiectivele (mai limitate ale acestei cărți). Oricum, ceea ce putem reține este că o bună conversație permite aprofundarea unei teme filosofice, tot așa cum rezolvarea de probleme sau experimentele de laborator permit, în știință, adâncirea înțelegerii materialului "teoretic".

Alegeți o temă și strângeți un grup interesat de discutarea ei! Acesta ar fi primul pas. Dar nu vă mulțumiți cu atât! Lipsa unor reguli ale dezbaterii va conduce rapid la eșec:

fie grupul care dezbate se va fragmenta (vorbesc mai mulți participanți deodată), fie nu va exista o focalizare a discuției și așa mai departe. Discuțiile filosofice sunt delicate și pot fi ușor ratate.

Prima regulă a și fost formulată: alegeți o temă! Am să iau drept exemplu discuția anterioară din carte. Tema ei ar putea fi rezumată astfel: mintea ca un subiect de reflecție independent.

Al doilea pas este alegerea unui text, pe care-l vor citi toți participanții la dezbateră. Nu este bine ca textul să fie foarte lung. De pildă, primele două meditații metafizice ale lui Descartes ar putea fi un text potrivit.

Oricum, este important să desemnați un lider al discuției. Primul său rol este acela de stabili două-trei obiective ale intervenției sale de început. Obiectivele vor exprima pe scurt ceea ce anume vrea să facă liderul discuției. De pildă, mergând din nou pe firul din această lucrare, ar putea fi vorba: (1) despre metoda folosită de Descartes pentru a argumenta în favoarea delimitării minții ca temă de reflecție filosofică distinctă; (2) existența minții ca o entitate aparte; (3) distincția dintre minte și corp.

Al doilea rol al liderului discuției este acela de a schița (într-un sfert de oră!) felul în care pot fi atinse, după ea sau el, obiectivele fixate. Limita de timp este foarte importantă! O discuție rodnică nu poate dura mai mult de o oră și jumătate – două ore. După aceea, capacitatea participanților de a se concentra scade. Cel mult dezbateră are caracterul unui divertisment, al unui mod de a-ți petrece plăcut timpul, după această limită.

Regula următoare este să separați din grup un panel (cinci-șase persoane). Imediat ce liderul discuției a lansat provocarea sa, fiecare membru al panelului pune o întrebare (de un minut maximum). Liderul discuției răspunde în circa cinci minute. După aceea, membrii panelului dezvoltă câte o idee proprie în orizontul dezbaterii deja conturate.

Este de așteptat ca aproximativ o oră să se scurgă după toate aceste etape. Întrebările celor din panel ar fi bine să fie de forma "Ce înțelegi prin ...?"

Timpu rămas este folosit pentru a lăsa restului membrilor grupului răgazul de a spune "ce au înțeles din discuție și ce nu au înțeles". În final, liderul discuției va încerca să precizeze liniile de forță ale dialogului. Nu va formula însă aprecieri sau concluzii cu privire la problemele abordate. Lucrul acesta este iarăși foarte important într-o discuție filosofică, unde ceea ce se câștigă este în zona distincției dintre "ceea ce am înțeles și ceea ce nu am înțeles", distincție care are o geografie variată de la un individ la altul. Ceea ce se câștigă prin conversația filosofică nu sunt cunoștințe noi, ci noi moduri de a pune întrebări, noi moduri de a vedea distincții conceptuale sau poziții teoretice.

O ultimă regulă este foarte importantă. Regula aceasta spune că aveți nevoie de un arbitru al discuției. Liderul discuției nu poate fi și arbitru.

Arbitrului îi revine rolul de a limita timpul intervențiilor participanților, de a ține o listă cu cei care pun întrebări sau vor participa la dezbateri. De asemenea, ea sau el trebuie să împiedice conversațiile colaterale sau întreruperile. Ca regulă generală, intervențiile arbitrului privesc respectarea formei dezbaterii; aceste intervenții nu vizează conținutul. Arbitrul nu ia poziție în chestiuni privitoare la conținutul dezbaterii.

Pentru fluenta discuției, ar fi extrem de util ca arbitra sau arbitrul discuției să țină, de fapt, două liste cu cei care pun întrebări. Pe prima listă sunt trecuți cei din panel sau, în final, dintre spectatori care pun întrebări mai elaborate și după tiparul deja indicat. Pe a doua listă sunt trecuți cei care vor să pună întrebări foarte scurte, punctuale. Aceste întrebări se referă la ceea ce se discută chiar în acel moment. Cel mai bine ar fi ca ele să privească nelămuriri legate de termenii folosiți sau de referirile făcute.

Puteți desigur experimenta cu aceste reguli. Ați putea, de pildă, renunța la separarea între panel și spectatori, dacă sunteți mai versați. În nici un caz nu renunțați însă la arbitru! Ceea ce puteți descoperi, variind regulile, este tocmai valoarea lor, mai mare sau mai mică, pentru rodnicia dezbaterii. Dincolo de aceste aspecte legate de formă, vom descoperi cât de mult ne lipsesc discuțiile filosofice în cultura noastră dominată de conversația de cafea, polemici vulgare și speculații superficiale.

## Experimentul lui Ryle

Ryle s-a născut chiar în 1900. A fost profesor de filosofie la Oxford. Când am fost la Oxford am auzit tot felul de anecdote despre Ryle. Una privea, de pildă, cunoștințele sale muzicale (Ryle nu avea, ca mulți filosofi, de altfel, ureche muzicală). Alta se referea la activitatea sa, în timpul războiului, în contraspionajul englez. Se zice că Ryle se plângea, la un moment dat, că un polițist dintr-o localitate englezească de provincie l-a arestat pe un spion german. Polițistul era, evident, zelos. Cei din contraspionaj îl știau însă de multă vreme pe spion. De ce nu-l arestau? Pentru că, explica Ryle, contraspionajul englez îl umplea pe spionul german cu tot felul de știri false. Adevărate sau nu, aceste anecdote conturează în parte și profilul filosofic al lui Ryle, pasiunea lui pentru reflecțiile tăioase, neașteptate.

În 1949, Ryle a publicat cea mai cunoscută carte a sa, *The Concept of Mind*.<sup>31</sup> Ca și în multe alte scrieri de filosofia minții și în *The Concept of Mind* experimentele gândite joacă un rol foarte important. Experimentul imaginat de către Ryle are drept subiect tocmai un vizitator, venit ca și mine cu dorința de a înțelege ce anume este Universitatea din Oxford. Străinului îi sunt arătate colegiile, bibliotecile, laboratoarele. I se explică felul în care funcționează sistemul de meditații (faimoasele *tutorials*, specifice Universităților din Oxford și Cambridge). Asistă la cursuri. Află câte ceva despre administrație.

Esențialul în experimentul lui Ryle este să ne închipuim toate acțiunile de mai sus așa cum s-ar desfășura ele, dacă străinul ar fi o persoană reală.

---

<sup>31</sup> Am folosit Gilbert Ryle, *The Concept of Mind*, Harmondsworth, Penguin Books, 1970, 316 p. O prezentare mai amplă a cărții este făcută în volumul *Dezbateri pe marginea problemei corp-minte*, editat de către mine în 1984. Exemplare din volumul respectiv sunt probabil greu de găsit, dar o variantă electronică va fi accesibilă în pagina mea de web (<http://www.fil.unibuc.ro/~solcan>). Prezentarea din 1984 este făcută pe baza traducerii lui Kurt Baier, *Der Begriff des Geistes*, Stuttgart, Philipp Reclam, 1978, 464p.

După aceste preliminarii urmează întrebarea cheie pusă de către străin: unde este însă Universitatea?

Ne putem, de asemenea, închipui o rugămintă nepotrivită: nu ar putea fi totuși să vadă Universitatea?

Asemenea chestiuni ridicate de către vizitator pun în evidență greșeala străinului: nu se poate vorbi separat despre Universitate și despre Colegiul Christ Church, despre Biblioteca Bodleiană, despre Ahmoleian Museum, ca și cum Universitatea din Oxford ar fi un alt membru al clasei obiectelor menționate.

### Tipurile și ideea de greșeală categorială

Dacă ne gândim puțin, ne dăm repede seama că ideea lui Ryle este simplă. O folosim adesea. Cu toții suntem în stare să distingem între clădirea unui colegiu și colegiu ca instituție. Clădirea are proprietatea de a fi vizibilă sau colorată ș.a.m.d. În nici un caz colegiul ca instituție nu este colorat (în sensul în care clădirea este colorată).

Tot așa, dacă privim jocul unei echipe de fotbal și ajungem să înțelegem ce rol are portarul sau cum merg pasele, nu ne vom apuca să întrebăm cui îi revine rolul de a furniza spiritul de echipă.

Obiectele (luate în sensul cel mai general, filosofic al termenului) aparțin unor clase sau categorii. Fiecare obiect dintr-o categorie are anumite proprietăți. Unele obiecte sunt, de exemplu, vizibile sau colorate.

Numerele naturale nu sunt însă colorate. Nu au acest gen de proprietate, tot așa cum clădirile colegiilor nu sunt pare sau impare. Dacă aș încerca să aflu care sunt numerele naturale roșii, cercetarea mea nu ar avea sens. Ea ar implica o greșeală categorială în sensul lui Ryle.

Vom vedea mai târziu că distincția aceasta între diverse tipuri este utilă nu numai în filosofia teoretică. Îndeletniciri mult

mai pragmatice, precum programarea computerelor, o folosesc din plin.

## Fantoma din corp

După Ryle, cartezienii comit o greșeală categorială atunci când vorbesc despre corp, care funcționează asemenea unei mașinării fizice, și despre minte, care ar avea rolul de a dirija corpul, dar ar avea o natură spirituală. Nu se poate spune despre minte că există ca o entitate separată.

Brațul meu, de pildă, are un anumit rol în mașinăria corpului. Tot așa degetele mele de la mână. Brațul va sluji la apropierea degetelor de tastele calculatorului, iar acestea vor apăsa pe taste și vor produce schimbări pe ecran, care mai târziu pot fi transformate în șiruri de semne într-o carte. Ar fi însă o greșeală categorială, după Ryle, să ne închipuim mintea ca fiind asemenea brațului, atâta doar că ea ar avea un rol în dirijarea mai subtilă a mașinării corpului.

După cum spune Ryle, nu ne-am luat o mănușă pentru mâna stângă, una pentru dreapta și, ca și cum ar fi ceva separat de acestea, o pereche de mănuși. Tot așa, nu putem vorbi despre corp și despre minte.

Contrapunerea dintre corp și minte, după Ryle, trebuie dizolvată. Ea nu are sens.

## Despre minte ... în limbajul obișnuit

*The Concept of Mind* are, ca să spunem așa, doi poli. Primul capitol respinge ceea ce Ryle numea dogma fantomei din mașină. Aici este dezvoltată ideea de greșeală categorială și argumentul împotriva minții și a mentalului ca entitate sau stări lăuntrice. Nu are sens să vorbim despre viața lăuntrică, care ar exista separat de corp și l-ar dirija. Acesta este primul pol al

argumentelor cărții. Ultimul capitol, consacrat psihologiei, constituie cel de-al doilea pol.

Ryle vede în evoluția psihologiei din prima jumătate a veacului al XX-lea o confirmare a ideii că "viața lăuntrică", neobservabilă nu poate fi subiectul unei investigații științifice. Nu există două lumi: o lume fizică și una spirituală, tot așa cum nu există mănua stângă și mănua dreaptă, pe de o parte, și perechea de mânuși, pe de alta. Desigur, putem vorbi despre perechea de mânuși. În limbajul obișnuit nu o transformăm însă într-o fantomă ontologică.

Psihologia de factură experimentală, consideră Ryle, nu are ce face cu fantoma din mașină. Ceea ce poate fi observat este, de fapt, comportamentul.

Filosofia minții, la Ryle, nu este totuși o analiză a fundamentelor psihologiei experimentale. De aceea pare a fi mai potrivit să vorbim despre această analiză ca despre un pol al cărții, polul opus cartezianismului. Ryle oferă, de fapt, între acești doi poli, o analiză a limbajului obișnuit menită să ne convingă că exprimările care se referă la o viață lăuntrică, privată sunt rodul unei concepții filosofice (greșite) și nu au nici un rol în limbajul uzual.

## **Limbajul comportamentului**

Înainte de a trece mai departe, trebuie să reamintesc că această carte este doar o introducere și să le sugerez celor care vor să se profesionalizeze în domeniul filosofiei minții (sau al filosofiei în general!) să mai înghită o pilulă amară. Dacă vreți să lucrați profesional trebuie să consultați textele filosofice în original. Desigur, aici nu putem oferi decât mici fragmente (cum am făcut-o deja). Ele sunt urmate, de altfel, de o parafrază în limba română a textului citat, astfel că puteți foarte bine sări peste fragmentele în limbi străine. Atâta doar că, atunci când vreți să știți mai mult, trebuie să vă aventurați în apele adânci ale textului original.



Filosofii fac distincții care nu sunt lesne de urmărit. Din punctul nostru de vedere, cred că momentul cel mai interesant la Ryle este legat de viziunea lui despre felul în care putem vorbi cu sens despre minte.

Ryle nu acceptă proiecțiile ontologice carteziene, dar acceptă ideea că putem vorbi despre minte. Cum facem acest lucru? Ryle crede că filosofia nu are nimic să ne spună în mod special despre minte. Oamenii știu deja să vorbească despre minte. Ei știu să utilizeze limbajul natural. Problema este că nu știu să vorbească despre conceptele mentale. Rolul filosofiei este să traseze harta conceptelor utilizate în limba obișnuită. În sensul acesta, filosofia este o cartografie conceptuală.

Din cele spuse până acum pare a rezulta că Ryle spune că a vorbi despre minte înseamnă a vorbi despre comportamente. Nu a abandonat oare psihologia experimentală referirile la o lume lăuntrică? Nu se mențin oare doar referirile la ceea ce poate fi observat, la comportament, cu alte cuvinte? Nu este chiar așa. Pe de o parte, Ryle nu este interesat de chestiunile ce țin de metodologia psihologiei. Pe de altă parte, el ne propune o abordare rafinată a limbajului comportamentului. Amintiți-vă de povestea cu spionul german! Iată ce scrie Ryle:

*When a cow is said to be a ruminant, or a man is said to be a cigarette-smoker, it is not being said that the cow is ruminating now or that the man is smoking a cigarette now. To be a ruminant is to tend to ruminate from time to time, and to be a cigarette-smoker is to be in the habit of smoking cigarettes.*

*The tendency to ruminate and the habit of cigarette-smoking could not exist, unless there are such processes as ruminating and smoking cigarettes.<sup>32</sup>*

Ryle distinge între procese (cum ar fi cel de rumegare, în cazul vacilor, sau de fumare a unei țigări, în cazul oamenilor) și capacitatea de a face ceva sau tendința, obiceiul de a se

<sup>32</sup> Gilbert Ryle, *The Concept Mind*, p. 143. Citatul este de la începutul celui de al doilea paragraf al Capitolului V (intitulat "Dispositions and occurrences").

comporta într-un anume fel (capacitate sau tendință despre care vorbim atunci când spunem că vaca este rumegătoare sau omul fumător).

Prin asemenea exemple, care pot părea cam fruste, Ryle introduce o distincție foarte importantă: cea dintre procesele comportamentale și capacitatea de a avea un anumit comportament. Când vorbim despre minte vorbim nu doar despre procese comportamentale, ci și despre capacități comportamentale.

### O discuție despre Gilbert Ryle

O discuție despre Ryle ar putea avea drept temă "fantoma din mașină" sau (de ce nu?) limbajul obișnuit. Cum vorbim (ca oameni obișnuiți) despre minte?

Dacă sunteți familiarizate sau familiarizați cu literatura filosofică și aveți o idee despre conceptele dispoziționale, atunci puteți discuta pe tema conceptelor (dispoziționale, în viziunea lui Ryle) cu ajutorul cărora vorbim despre capacitatea sau tendința de a ne comporta într-un anume fel.

Nu aș recomanda însă excesul de discuții pe teme care presupun un avantaj amplu de lecturi filosofice. După mine, focalizarea discuțiilor asupra textelor de bază ca atare este foarte importantă pentru formarea deprinderilor de a dezbate filosofie. Altfel, toată conversația riscă să se prăbușească sub presiunea recursului la argumentul autorității (cutare autoritate a spus că intenția lui Ryle sau înțelesul textului lui Ryle este ...!).

Din acest motiv nici nu voi mai menționa expres îndemnul de a organiza discuții pe o temă sau alta. Temele se nasc din lectura textului însuși. Restul este o chestiune ce ține de disciplină (de respectarea unor reguli) ale discuției.

De pildă, puteți investiga diversele aspecte ale distincției lui Ryle între obiceiuri și capacități inteligente.<sup>33</sup>

---

<sup>33</sup> Gilbert Ryle, *The Concept of Mind*, p.42 ș.u.

## Cum descoperiți o temă interesantă?

Probabil, răspunsul cel mai simplu la întrebarea din titlu este că nu pot fi formulate reguli pentru descoperirea de teme interesante. În parte, așa și este. Sunt multe motive pentru care ceva poate fi interesant sau neinteresant.

Pentru filosofie există însă o provocare în întrebarea de mai sus. Să ne gândim puțin la ceea ce face Ryle! El a identificat un mit sau o legendă privitoare la existența a două lumi, una corporală, fizică, alta spirituală, lăuntrică. Această legendă este foarte răspândită. Ea stă la originea multor presupoziții subiacente discursului referitor la minte.

Ce sunt presupozițiile? Ideea de presupoziție, deși permite dezvoltări filosofice subtile, este ușor de înțeles pe baza unui exemplu deja discutat, cel referitor la proprietățile numerelor naturale.

Un număr natural  $n$  poate fi par sau impar. Se presupune deci că numărul  $n$  poate fi împărțit la doi, cu sau fără rest. Presupoziția aceasta este întotdeauna adevărată despre un număr natural. Enunțul " $n$  este un număr par" este însă adevărat sau fals, după cum restul la împărțirea cu doi este sau, respectiv, nu este nul.

În schimb, presupoziția că numerele naturale au o culoare este falsă. De aceea nici nu mai are sens să cercetăm dacă numărul  $n$  este sau nu roșu.

Tot așa, am putea spune că, în cazul formelor geometrice, proprietăți precum "pătrat" sau "rotund" se exclud una pe alta. O formă este sau rotundă sau pătrată. A căuta pătratul rotund este lipsit de sens. Dacă totuși vrem să folosim cuvintele "pătrat" și "rotund" pentru proprietăți care revin aceleiași forme geometrice, atunci trebuie să admitem că am schimbat presupoziția și că am conferit cuvintelor respective un sens nou.

Acum, ceea ce fac de multe ori filosofii este să identifice presupoziții sau, mai exact, sisteme de presupoziții. Acest lucru îl face și Ryle când conturează ceea ce el numește "legenda celor două lumi". Când aceste presupoziții au efecte asupra

unor enunțuri care vi se păreau poate inocente, atunci puteți identifica teme pe care o discuție este cu adevărat interesantă.

De pildă, toți oamenii care au ceva cultură și credința că pot vorbi pe teme filosofice ar fi tentați să întrebe cum stau lucrurile cu actele de voință. Care ar fi statutul sau definiția voinței?

Ei bine, după Ryle, volițiile sau actele de voință presupun admiterea legendei celor două lumi. Să tragi mânerul semnalului de alarmă din tren este, dacă presupunem că există o lume fizică și una lăuntrică, un act fizic care are drept cauză un act mental, cel de voință.<sup>34</sup>

Împărțirea tripartită a minții sau sufletului în gândire, sensibilitate și voință nu are sens dacă presupoziția existenței unei lumi lăuntrice este falsă. Se vede de aici că, de fapt, câștigul obținut atunci când învățăm filosofie este unul în planul griii cu care folosim conceptele. Cei neantrenați în chestiunile filosofice vor utiliza cu nonșalanță idei precum cea de act de voință.

Filosofi precum Ryle nu te pot împiedica, la urma urmei, să vorbești despre acte de voință sau orice altceva. Dar nu mai poți face acest lucru ca și cum ar fi ceva neproblematic sau care ar avea presupozii care trebuie admise de toată lumea.

Poți vorbi și despre pătrate rotunde, dar trebuie să oferi o explicație filosofică a felului în care utilizezi aceste concepte și a motivelor pentru care ai adoptat o presupoziție sau alta.

În rezumat, pentru a avea o discuție mai interesantă încercați să identificați presupozii larg acceptate care ar putea să cadă sub briciul analizei filosofice. Aerul paradoxal al concluziilor va da un farmec aparte dezbaterii și, în orice caz, veți deveni mai conștiente sau mai conștienți de capcanele care se ascund în spatele unor cuvinte ce altfel pot părea absolut nevinovate.

---

<sup>34</sup> Vezi, *mutatis mutandis*, exemplul analog al lui Ryle în *The Concept of Mind*, p.62 (începutul celui de al doilea paragraf al capitolului al treilea, dedicat voinței).

## De la Descartes la Ryle

Din când în când este bine să vedem ce obiective a avut demersul nostru și cum le-am atins. Am început discuția despre analiza filosofică a minții cu Descartes și am identificat următoarele repere:

- **Experimentul cu geniul cel rău**
- **Mintea ca subiect de reflecție independent**
- **Dualitatea corp-minte.**

Mintea este un subiect de reflecție independentă și pentru Ryle, dar el respinge statutul ontologic pe care-l conferă minții Descartes. Dispare astfel și toată problema modului în care stările mentale pot fi cauze ale actelor fizice. În cazul lui Ryle, reperele de bază sunt:

- **Distingerea tipurilor**

*{greșelile categoriale ne forțează să distingem între diferite tipuri de obiecte și să fim atenți ce proprietăți se presupune că au sens pentru obiectele de un anume tip}*

- **Respingerea dualismului**

*[Ryle respinge ceea ce el numește fantoma din mașină. A vorbi despre minte nu este decât un alt fel de a vorbi despre comportamente.]*

- **Distincția dintre procese și capacități comportamentale**

*[Ryle distinge între comportamente așa cum se desfășoară ele ca procese actuale și capacitatea de a te comporta într-un anumit fel.]*

Poziția filosofică a lui Ryle se numește *behaviorism logic*. Behavior înseamnă în limba engleză "comportament". Ryle argumentează cu ajutorul analizei logice a limbajului în favoarea behaviorismului.

## Ludwig Wittgenstein și ideea tipurilor de expresii

Toți cei care au deschis această carte, chiar dacă nu și-au pus problema de mai sus, au auzit probabil de Ludwig Wittgenstein. Născut în 1889 la Viena, el a studiat la Cambridge între 1911 și 1914. Ideile care ne interesează aici au fost descoperite de mintea originală a lui Wittgenstein în perioada când revine la Universitatea din Cambridge între 1929 și 1947. Ele au fost publicate însă abia după moartea sa survenită în 1951.<sup>35</sup>

Conturarea ideilor filosofice ale lui Wittgenstein depășește cu mult obiectivele pe care ni le-am propus în această carte. Fixarea câtorva puncte de reper prin *Philosophical Investigations* nu ar fi însă, cred eu, rău venită înainte de discutarea problemei care ne interesează în mod special aici.

Wittgenstein, ca și mulți alți filosofi, este interesat nu de ceea ce este, ci de modurile în care ceva poate fi:

*Es ist uns, als müßten wir die Erscheinungen durchschauen: unsere Untersuchung aber richtet sich nicht auf die Erscheinungen, sondern, wie man sagen könnte, auf die „Möglichkeiten“ der Erscheinungen.*<sup>36</sup>

Cercetarea filosofică nu încearcă să pătrundă fenomenele sau “natura” acestora. Ea este îndreptată către posibilități. De aici și interesul lui Wittgenstein pentru tipurile de enunțuri.

Poate chiar mai mult decât la Descartes sau Ryle, experimentele imaginare joacă un rol deosebit la Wittgenstein.

<sup>35</sup> În *Philosophical Investigations*. V. Ludwig Wittgenstein, *Philosophical Investigations*, text german și traducere în limba engleză, Oxford, Basil Blackwell, 1953, 232 p. (duble; textul german este pus pe paginile numerotate normal, cel englez pe cele al căror număr este urmat de litera “e”).

<sup>36</sup> Ludwig Wittgenstein, *Philosophical Investigations* I, aforismul 90. Cartea este, de fapt, o colecție de mici texte, aforisme. Referirea la ele facilitează consultarea altor ediții decât cea din 1953 sau a unor antologii.

Să zicem că suntem familiarizați cu un tip de enunțuri pe care Wittgenstein le numește “obiect și denumire”. De pildă, înțelegeți ce înseamnă “Am o carte”, “Am o floare” și așa mai departe. În fiecare caz este vorba de posedarea unui obiect desemnat prin cuvintele “carte”, “floare” etc. Sunt oare aceste enunțuri de același tip cu “Am o durere”?

Wittgenstein ne propune un experiment imaginar pentru a ne arăta diferența de tip care intervine mai sus. Să ne închipuim că fiecare are o cutiuță în care ține un gândac pe care nu-l arată nimănui altcuiva.<sup>37</sup> Nimeni nu a văzut altceva decât gândacul din cutia sa. Atunci când folosesc cuvântul “gândac” oamenii sunt obligați să se ghideze doar după ceea ce au văzut în cutia lor. Cutia ar putea fi și goală. Gândacii ar putea fi foarte diferit colorați.<sup>38</sup> A folosi cuvântul “gândac” ca și cum ar fi o denumire pentru un obiect, ca mai sus, este o eroare.

*Das heißt: Wenn man die Grammatik des Ausdrucks der Empfindung nach dem Muster von „Gegenstand und Bezeichnung“ konstruiert, dann fällt der Gegenstand als irrelevant aus der Betrachtung heraus.<sup>39</sup>*

Obiectul devine irelevant, în condițiile din experimentul imaginat de către Wittgenstein, în expresia “Am un gândac”. Tot așa stau însă lucrurile și în cazul, foarte important pentru filosofia minții, al expresiei “Am o durere”!

## Problema limbajului privat

Unde mergem mai departe pe drumurile filosofiei? Mai sus am ajuns deja la ideea că expresiile în care este vorba despre senzații (cea de durere, de exemplu) nu sunt de același tip cu expresiile referitoare la obiecte accesibile public. Nu aş putea însă să-mi fac reguli private, reguli știute doar de mine și pe care

<sup>37</sup> Atenție la analogia cu citirea unui text fără a rosti ceva. Fiecare ține pentru sine ceea ce are.

<sup>38</sup> Atenție din la faptul că, fără puțința de a călători prin lumea variatelor posibilități, experimentul filosofic nu ar putea avea loc.

<sup>39</sup> Ludwig Wittgenstein, *Philosophical Investigations* I, 293.

doar eu să mă pricep să le urmez? Aceste reguli ar putea fi ascunse de privirile curioase ale celorlalți chiar și dacă nu există o lume lăuntrică. În fond, oamenii citeau vreme de secole cu voce tare și doar relativ târziu au învățat să ascundă comportamentul verbal și să parcurgă textele în tăcere. Astfel aș putea vorbi privat despre gândacul din cutie.

Wittgenstein vorbește despre un paradox ce se naște atunci când încercăm să urmăm reguli:

*eine Regel könnte keine Handlungsweise bestimmen, da jede Handlungsweise mit der Regel in Übereinstimmung zu bringen sei.<sup>40</sup> (...) Darum ist „der Regel folgen“ eine Praxis. Und der Regel zu Folgen **glauben** ist nicht: der Regel folgen. Und darum kann man nicht der Regel „privatim“ zu folgen, weil sonst der Regel zu folgen glauben dasselbe wäre, wie der Regel folgen.<sup>41</sup>*

Pasajele acestea cam sibilinice au primit o interpretare celebră din partea lui Saul Kripke. Filosoful american a oferit un exemplu simplu, menit să ne arate de ce nu putem stabili ce regulă urmăm atunci când facem ceva.<sup>42</sup>

Să spunem că m-am închis într-o cameră și am un set de fișe. Scriu pe câte două fișe numere naturale. Apoi iau o altă fișă și trec pe ea un număr. Rezultatele ar putea fi puse pe trei coloane: pe primele două sunt numerele de pe primele două fișe, iar pe a treia cele de pe fișa suplimentară:

4	4	8
16	4	20
53	7	60

S-ar părea că adun primele două numere, în sensul cel mai obișnuit. Dar de ce nu ar fi vorba despre ceea ce am putea numi o semiadunare? Aceasta ar da aceleași rezultate cu

<sup>40</sup> Ludwig Wittgenstein, *Philosophical Investigations* I, 201.

<sup>41</sup> Ludwig Wittgenstein, *Philosophical Investigations* I, 202.

<sup>42</sup> Cf. Saul Kripke, *Wittgenstein on Rules and Private Language*, Oxford, Basil Blackwell, 1982, p.9.



adunarea până ce s-ar atinge un "număr magic". De acolo înainte rezultatele ar fi altele:

299     1     300

300     1     7

Aici numărul magic este 300. După atingerea numărului magic rezultatul este constant 7. De ce aş trage aşadar concluzia că urmez regula adunării şi nu pe cea a semiadunării?

Numere magice de genul celui de mai sus pot fi însă imaginate la nesfârşit. Iar aceasta nu înseamnă altceva decât că nu pot privilegia în mod lipsit de ambiguitate o regulă sau alta. A urma o regulă este, după Wittgenstein, o practică şi o chestiune de interpretare.

### După Wittgenstein

Implicaţiile acestei filosofii wittgensteiniene trec mult dincolo de filosofia minţii. După trecerea lui Wittgenstein prin lume însă nu mai putem discuta în mod necritic despre senzaţii private sau un limbaj privat. Cuvântul "necritic" este cheia. După cum vom vedea mai jos, există în filosofia actuală a minţii, discuţii despre un *limbaj privat*. Ipoteza existenţei sale este însă rodul unor reflecţii foarte elaborate şi nu se bazează pe acceptarea necritică a unor susţineri naive. De altfel, pe această linie s-ar situa şi rezumatul întâlnirii noastre aici cu Wittgenstein:

- **Experimentul cu gândacul din cutie**

*{Experimentul gândit propus de Wittgenstein forţează distingerea expresiilor de tipul "Am o carte" de cele de tipul "Am o durere".}*

- **Argumentul Wittgenstein-Kripke**

*{Respingerea ideii că regula urmată atunci când acţionăm poate fi identificată în chip univoc. Kripke i-a dat o formă mai operaţională prin indicarea unui mod de a construi contraexemple.}*

Wittgenstein ajunge, folosind metodele indicate mai sus, la

- **Respingerea limbajului privat**

*[Regulile nu pot fi urmate în mod privat. Nu poate exista un limbaj privat.]*

Filosofia lui Wittgenstein este aproape imposibil de clasificat. În orice caz, Wittgenstein nu ar fi acceptat ideea că filosofia sa a minții este o variantă a behaviorismului. Cert este însă că Wittgenstein reprezintă un reper extrem de important în filosofia secolului al XX-lea. De aici și titlul acestui paragraf. Filosofia arată altfel după trecerea lui Wittgenstein prin lume.

Nodul cel mai important mai sus este problema limbajului privat. După cum am spus, există filosofi care s-au încumetat să meargă pe alt drum decât Wittgenstein. Dar, pentru a le pricepe gândurile, trebuie să deprindem mânuirea unor noi unelte conceptuale.

## **Problema raportului corp-minte**

Nu sunt puține scrierile care aproape că identifică filosofia minții cu discutarea problemei raportului corp-minte. Pozițiile în această dezbatere sunt atât de numeroase încât fie și prezentarea rezumativă a celor mai importante dintre ele ar umple rapid spațiul rezervat pentru această carte.

Ceea ce ne interesează aici este că, o dată cu Descartes, problema raportului corp-minte se conturează ca o problemă distinctă de probleme metafizice mai vechi, precum cea a raportului dintre materie și spirit. De altfel, sunt posibile variate combinații ale pozițiilor în problemele amintite.

Poziția lui Descartes în problema corp-minte era dualistă. În veacul al XX-lea cele mai populare au fost însă orientările antidualiste de un fel sau altul. Am prezentat deja argumentele antidualiste ale lui Ryle.

Cea mai cunoscută dintre orientările antidualiste este cea fizicistă. Ea își are originea în proiectul Cercului de la Viena de a reduce toate științele empirice la fizică. În acest context,

psihologia urma și ea să fie redusă la fizică.<sup>43</sup> Din punct de vedere filosofic cea mai semnificativă consecință a fizicalismului este ideea că o cunoaștere completă a proprietăților fizice ale creierului este suficientă pentru a explica orice stare mentală.

Ceea ce pare a uni variatele orientări antidualiste este ideea că organismele sunt mecanisme fizico-chimice.<sup>44</sup>

Cum ar putea răspunde însă fizicaliștii la obiecții de genul “senzațiile sunt private, dar procesele din creier sunt publice”? După J.J.C.Smart, când spunem că am avut cutare sau cutare senzație folosim limbajul introspecției.<sup>45</sup> Limbajul introspecției are o logică diferită de cea a limbajului proceselor materiale. Dar, atunci când vom avea o teorie îmbunătățită a

---

<sup>43</sup> Cercul de la Viena respingea vitalismul în biologie și psihologie. În Manifestul lor, membrii Cercului scriau: “Formele lingvistice pe care le mai folosim astăzi în psihologie își au originea în anumite noțiuni metafizice antice despre suflet. (...) [în domeniul psihologiei – n.n.] aproape tot ceea ce privește analiza epistemologică urmează să fie făcut, bineînțeles că aici analiza este mult mai dificilă decât în fizică. Încercarea psihologiei behavioriste de a sesiza psihicul prin comportarea trupurilor, prin ceea ce se găsește la un nivel accesibil percepției este prin poziția sa principală aproape de concepția științifică despre lume.” (“Cercul de la Viena [manifest]”, traducere Eva Galambos, în *Filosofia analitică*, volum redactat de Mihail Radu Solcan, București, Academia “Ștefan Gheorghiu” și Universitatea din București, 1982, p.69). Proiectul Cercului a fost să editeze o *Enciclopedie internațională a științei unificate*. Problema pe care și-o puneau membrii Cercului era, în fond, aceea a unității de limbaj a științei (v. *Filosofia analitică*, p.79). Orice enunțuri urmau a fi reduse la enunțuri formulate într-un sublimbaj fizical. Acest limbaj urma să cuprindă doar termeni care se referă la obiecte și procese fizice, la proprietăți observabile ale acestora și predicate dispoziționale (de genul “solubil”, “elastic” etc.). Această reducere la limbajul fizical trebuia să funcționeze și în cazul psihologiei. De exemplu, “furios”, spunea Rudolf Carnap, poate fi definit în termenii comportamentului observabil al unei persoane și al dispoziției acesteia pentru anumite reacții (cf. *Filosofia analitică*, p.81).

<sup>44</sup> V., de pildă, în acest sens, remarcile lui J.J.C.Smart, “Sensations and Brain Processes”, în David M. Rosenthal (editor), *The Nature of Mind*, Oxford University Press, 1991, p.169. Dacă reconstituim “arborele bibliografic”, adică dacă mergeți pe firul citărilor din textul lui Smart, atunci veți da peste numele unor autori precum Paul Oppenheim sau Hilary Putnam (articol din 1958 despre unitatea științei), Feigl, Place.

<sup>45</sup> J.J.C.Smart, *Op.cit.*, pp.173-174.

proceselor din creier, vom putea spune că prezența unui proces cerebral sau a altuia înseamnă că subiectul are o anumită senzație. Până atunci nu dispunem de nimic altceva decât de exprimările în limbaj introspectiv.

Cum ar putea fi însă contracarat experimentul cu gândacul din cutie imaginat de Wittgenstein, dacă acest experiment ar fi să fie folosit împotriva fizicalismului? Rațiunea folosirii acestui experiment împotriva fizicalismului este simplă: limbajul fizical este de genul “obiecte și denumiri”, cum spune Wittgenstein. Vedem ceva și acel ceva este un obiect și denumim obiectul respectiv “carte”.

Dacă analizăm răspunsul lui Smart,<sup>46</sup> ne dăm seama că el admite o diferență de tip între expresiile care relatează despre obiecte precum cartea și care raportează o senzație. Diferența este însă, spune Smart, de genul celei pe care o facem atunci când spunem că “acest obiect arată astfel” (fără nici un fel de rețineri) și “acest obiect arată precum ...”. Este ca și cum am spune “acest obiect arată precum o pajiște verde”. De aceea, spune Smart, nici nu există un limbaj al proprietăților private ale lucrurilor. Noi trebuie să ancorăm, ca să spun așa, expresiile de al doilea tip în situații care permit relatări de primul tip. Dar toate aceste procese mentale, inclusiv cel de ancorare, sunt procese cerebrale.

## Funcționalism și complexitate

În câte feluri puteți face ca o ușă să nu poată fi deschisă pentru o vreme? Puteți evident să încuiăți ușa cu cheia. Puteți trage un zăvor. Puteți pune un scaun în ușă. Scaunul are, în acest caz, aceeași funcție cu zăvorul. Putem vorbi despre o proprietate funcțională; proprietatea respectivă o are și zăvorul, o are și scaunul – multe lucruri pot avea aceeași funcție.

Chiar dacă un avantaj al reducerii proceselor mentale la procese cerebrale, față de reducerea la comportamente

---

<sup>46</sup> J.J.C.Smart, *Op.cit.*, p.174.

observabile externe, este acela că pot exista procese cerebrale la care acces are doar persoana în al cărei creier se petrec, fără ca să existe un comportament observabil extern, există totuși o dificultate a acestui tip de teorii. De ce ar fi o senzație, de pildă, identificabilă cu un anume proces din creier? Ar fi ca și cum ne-am închipui că ușa se poate închide doar cu zăvorul!

Soluția la acest tip de problemă este funcționalismul: teoria minții conform căreia ceea ce contează este funcția pe care o îndeplinește un proces fizic. Procese diferite pot avea aceeași funcție. Atenția se concentrează acum asupra sistemului de proprietăți funcționale. Acestea ar putea fi realizate și de altceva decât neuronii din creier.

Unul dintre pionierii înțelegerii în acest fel a minții a fost Hilary Putnam.<sup>47</sup> El a subliniat faptul că explicația funcționalistă a minții oferă posibilitatea ca procesele mentale să se desfășoare nu doar în creier, ci pe orice alt fel de suport convenabil (în sensul că permite implementarea funcțiilor relevante).

Mai târziu, Hilary Putnam a dezvoltat o serie de considerații care trec dincolo de ideea reducerii minții la un sistem de proprietăți funcționale ale elementelor de bază (sau, pe scurt, ale microstructurii). Când ansamblul devine foarte complex, legile proceselor de la nivelul microstructurii nu au neapărat relevanță pentru macrostructură.

Hilary Putnam oferă exemplul simplu în fond al evoluției biologice. Evoluția depinde de microstructură (variații ale genotipului), dar mai depinde și de condiții (precum prezența oxigenului) care sunt accidentale din punctul de vedere al legilor fizicii.<sup>48</sup>

---

<sup>47</sup> Vezi, de exemplu, studiul său publicat inițial în 1967 "Psychological Predicates", antologat cu titlul "The Nature of Mental States" în David M. Rosenthal (editor), *The Nature of Mind*, Oxford University Press, 1991, pp.197-203. Putnam a început să critice funcționalismul în a doua jumătate a anilor 1970. Pentru mai multe repere ale operei lui Putnam vezi capitolul "În jungla bibliografiei".

<sup>48</sup> Cf. Hilary Putnam, "Reductionism and the Nature of Psychology", în John Haugeland (editor), *Mind Design*, MIT Press, 1981, p.209.

Cu alte cuvinte, folosirea scaunului pentru a închide ușa este de nededus din proprietățile sale fizice. Ar putea fi, de altfel, multe alte funcții pe care să le joace scaunul. Iar când intervine în joc complexitatea nu are nici un rost să ne îndreptăm atenția către microstructură. Un element sau altul al acesteia, din punctul de vedere al legilor microstructurii, poate să aibă o natură pur accidentală. Funcția pe care o îndeplinește poate fi însă una cheie la nivelul macrostructurii.

Când complexitatea este introdusă în acest fel în jocul argumentelor, ideea reducerii psihologiei la fizică, unul dintre pilonii fizicalismului, devine implauzibilă.

## Experimentul lui Frank Jackson

Probabil că v-ați obișnuit deja cu ideea că experimentele gândite sunt frecvent utilizate în filosofia minții. Ele au un rol extrem de important în critica teoriilor filosofice. Departe de a fi ceea ce cred unii, filosofia nu oferă posibilitatea de a spune orice oricum. Poți spune orice, dar nu oricum. O dată puse în evidență defectele unei teorii, aceasta trebuie cel puțin schimbată. De asemenea, probleme nebănuite ies la iveală sau altele mai vechi capătă noi semnificații.

Aceasta a fost și soarta fizicalismului. Frank Jackson a arătat cum putem pune în dificultate fizicalismul dacă respectiva teorie este înțeleasă în sensul că o descriere completă din punctul de vedere al proprietăților fizice ne oferă o cunoaștere completă. În particular, o cunoaștere completă din punct de vedere fizic a creierului ar fi o cunoaștere completă.

Să vedem acum care este experimentul lui Frank Jackson.<sup>49</sup> Personajul din experimentul imaginar, Mary, este o fată care trăiește izolată într-o cameră. În cameră nu există obiecte colorate, iar contactul cu lumea este făcut prin intermediul unor ecrane alb-negru. În schimb, Mary are

---

<sup>49</sup> Expus aici după Frank Jackson, "What Mary Didn't Know", în Rosenthal (ed.), *The Nature of Mind*, pp.392-394.

cunoștințe strălucite de fizică. Ea știe tot ce se poate ști despre proprietățile fizice ale obiectelor.

De exemplu, Mary știe totul despre proprietățile fizice ale roșiilor, deși nu a văzut și nici nu a mâncat vreodată o roșie. Mary dispune însă de toate informațiile obținute prin analiza de laborator a roșiilor. Ea are, cu alte cuvinte, un tablou complet al proprietăților fizice ale roșiilor.

Într-o zi, Mary iese din cameră. Ea ajunge să vadă o roșie. Pentru prima dată are senzația de roșu, vede efectiv culoarea roșie și nu nuanțe de gri. Aceasta este, spune Frank Jackson, o extindere a cunoașterii de care dispunea Mary.

Argumentul lui Jackson este următorul: experimentul ne forțează să admitem că o cunoaștere completă a proprietăților fizice nu este și o cunoaștere completă în general. Mary a aflat ceva nou atunci când a ieșit din cameră. Ea aflat cum arată roșiile când le percepi și culoarea.

Proprietățile stărilor noastre mentale care survin atunci când avem o trăire de genul celei pe care o are Mary atunci când iese din cameră și vede culoarea roșiilor sunt denumite în filosofie *qualia*.<sup>50</sup>

### De la fizicalism la *qualia*

Cititoarea sau cititorul nu au aflat desigur, după parcurgerea acestui capitol, tot ce s-a spus în filosofia minții practică

<sup>50</sup> *Qualia* (termenul este la plural) înseamnă în limba latină "calități" (v. G. Guțu, *Dicționar latin-român*, București, Editura științifică și enciclopedică, 1983, p.1013). Folosindu-l la singular, Cicero spune *quale bellum nulla barbaria gessit* "război, precum lumea barbară n-a purtat vreodată" (apud Guțu, *Op.cit.*). Altfel spus, "război ca acesta nu a purtat vreodată lumea barbară". Dificultatea sintaxei lui "ca" poate fi apreciată de oricine a încercat să facă analiza logică a limbii naturale. Experimentul lui Jackson cred că pune bine în lumină resorturile mai profunde ale problemei. Ce înseamnă oare "a fi ca o roșie"? De unde am putea ști dacă este sau nu ceva ca o roșie? Știm oare doar pe baza măsurătorilor dintr-un laborator? Experimentul imaginar al lui Frank Jackson ne arată că, pentru a ști acest lucru, trebuie aduse în joc proprietățile calitative ale stărilor în care se află mintea noastră când vedem o roșie.

cu mijloace filosofice tradiționale. Nici nu a fost acesta obiectivul nostru. Sper, în schimb, ca ei să știe cum poate fi făcută filosofia minții.

Filosofia construiește și distruge distincții conceptuale, aduce la lumină presupuziții, folosește experimente imaginare. În special în filosofia minții se remarcă recursul la experimentele gândite. Toate acestea sunt punctele forte ale demersului filosofic. Slăbiciunile?

Una dintre slăbiciuni pare a fi ceea ce am putea denumi indecizia filosofiei. Dualism sau fizicalism? Discutați și veți vedea că fiecare parte are argumentele sale. Nici unul nu este însă hotărâtor. După mine, aceasta este însă o slăbiciune aparentă. Iar dacă privim lucrurile cu înțelepciune s-ar putea ca, în fapt, să fie aici o virtute crucială. Oamenii au nevoie, în activitățile cotidiene, în știință, în artă, de ipoteze alternative, între care să aleagă. Rolul discuției filosofice este să mențină în viață această diversitate, nu să o ucidă.

Posibilele slăbiciuni ale filosofiei rezidă în interfața ei cu alte preocupări umane. Dar despre aceasta vom vorbi imediat, în finalul capitolului. Să rezumăm deocamdată parcursul nostru între fizicalism și *qualia*.

- **Fizicalismul**

*[Sub acest nume ar putea fi reunite, la limită, variatele orientări antidualiste. În sens strict, fizicalism înseamnă însă reducerea legilor psihologiei la legile fizicii.]*

- **Procese mentale și procese cerebrale**

*[Avantajul identificării proceselor mentale cu procesele din creier este că devine inteligibilă existența unor procese care nu sunt însoțite de un comportament observabil extern.]*

- **Funcționalismul**

*[Avantajul recursului la sisteme de proprietăți funcționale este că putem înțelege procesele mentale independent de ceea ce se întâmplă să fie substratul lor fizic.]*

- **Rolul complexității**



*{Invocarea complexității este, în fapt, un recurs la contraexemple. Contraexemple relevante ne forțează să admitem că legile microstructurii fizice pot să nu ne spună nimic interesant despre macrostructură.}*

- **Experimentul lui Frank Jackson**

*{Experimentul imaginar propus de Frank Jackson urmărește să forțeze ideea că o cunoaștere a proprietăților fizice nu este o cunoaștere completă. În orice caz, ea nu ar da socoteală de proprietăți calitative ale minții.}*

- **Qualia**

*{Qualia este un termen convenabil pentru proprietățile sistemelor mentale care survin atunci când, de pildă, percepem culori.}*

Din punctul de vedere al fizicalismului, *qualia* sunt reductibile la proprietăți fizice (sau ar trebui să fie, pe măsură ce se dezvoltă știința). Din punctul de vedere al dualismului, *qualia* nu sunt reductibile la proprietăți fizice.

## Un divorț cu acordul ambelor părți<sup>51</sup>

De unde le vin filosofilor intuițiile cu privire la minte? Ideile le vin din lectura critică a scrierilor altora. Dar intuițiile? Intuițiile au drept sursă examinarea propriei minți, introspecția.

Nu are sens să vorbim de introspectarea altor minți. Acesta ar fi un nonsens. Putem, în schimb, examina comportamentul altora. Atunci baza nu doar a intuițiilor noastre se schimbă, ci și metodologia cercetării. Experimentele empirice devin posibile.

Acestea sunt considerente abstracte, dar ele redau liniile de forță ale unei situații care avea corespondent în realitate în

---

<sup>51</sup> Metafora divorțului și, mai târziu, cea a reînfripării mariajului sunt preluate de la Pascal Engel, *Introduction à la philosophie de l'esprit*, Paris, La Découverte, 1994, 256 p.

secolul al XIX-lea. Filosofia și psihologia continuau să fie strâns legate în acea vreme.

Legătura era bilaterală. Se spunea, în filosofie, că logica studiază "legile gândirii corecte". Spre sfârșitul veacului al XIX-lea și începutul secolului al XX-lea tendința de separare a devenit tot mai pronunțată. Logica a încetat să mai fie o cercetare a gândirii de vreun fel sau altul. Motivul era limpede: gândirea este cercetată empiric. Filosofia nu are nimic de a face cu experimentele empirice ale noii psihologii.

Două mari nume ale istoriei filosofiei au marcat acest divorț de psihologie: Frege și Husserl. Frege deplânge preponderența în filosofie "a metodelor psihologice de analiză, metode care au pătruns până și în domeniul logicii"<sup>52</sup> El face o distincție între

*reprezentări, pe de o parte, concepte și obiecte, pe de altă parte.*<sup>53</sup>

Frege scrie că

*trebuie să distingem în mod riguros psihologicul de logic, subiectivul de obiectiv.*<sup>54</sup>

Acesta este un principiu care călăuzește cercetările lui Frege. Logica modernă, ale cărei baze le pune Frege, pornește nu de la o investigație a legilor gândirii, ci a relațiilor de consecință logică. Terenul pe care se mișcă astfel nu este cel al gândurilor sau reprezentărilor, lăsate pe seama psihologiei, ci al construcțiilor și enunțurilor din limbaje ideale.

---

<sup>52</sup> Gottlob Frege, "Fundamentele aritmeticii", în *Scrieri logico-filosofice*, trad. de Sorin Vieru, București, Editura științifică și enciclopedică, 1977, p.33.

<sup>53</sup> Frege, *Op. cit.*, p. 37.

<sup>54</sup> *Ibidem*. Sorin Vieru, în comentariile sale la "Fundamentele Aritmeticii" arată că "Pentru Frege, logica reprezintă prin excelență o investigație a obiectivului, în timp ce psihologia poartă asupra subiectivului" (p.159).

Mai mult decât atât, contribuția lui Frege este crucială pentru începuturile filosofiei analitice. Iar, lucru și mai important, filosofia analitică, în special atunci când este văzută ca analiză logică a limbajului, este o filosofie antipsihologistă.

La Husserl regăsim aceeași pledoarie antipsihologistă. El reproșează psihologismului că transformă adevărurile logice din legi absolut exacte având un caracter conceptual în probabilități mai mult sau mai puțin vagi, bazate mai degrabă pe experiență și inducție.<sup>55</sup>

Frege și Husserl stau la originea celor mai influente tendințe filosofice ale secolului al XX-lea, filosofia analitică și fenomenologia. Nu este de mirare că divorțul dintre psihologie și filosofie a fost profund și durabil.

La rândul lor, psihologii au dorit despărțirea de filosofie. Pe de o parte, conform unei bine-cunoscute legende, desprinderea de filosofie marchează începuturile științelor naturii. Or, psihologia se vrea - cu deosebire către sfârșitul veacului al XIX-lea - o știință experimentală. Pe de altă parte, există rațiuni interne care împing psihologia către experiment și departe de filosofie.

Psihologia se orientează către observarea comportamentului. Acesta este singurul accesibil unui observator extern. Se ajunge la punerea între paranteze a mentalului. Este vremea când psihologia este dominată exclusiv de către *behaviorism*. Pentru behavioriști fenomenele mentale nu pot face obiectul studiului experimental și, până la urmă, orice referire la ele are statutul conceptelor naive, preștiințifice.

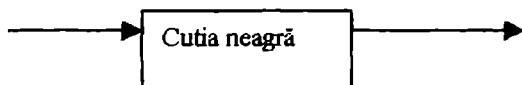
### Și totuși geniul cel rău?

O convenție terminologică privitoare la înțelesul lui "behaviorism" pare extrem de ușor de încheiat. Cu toții am auzit

---

<sup>55</sup> Cf. Edmund Husserl, *Logische Untersuchungen*, vol. I, Halle, Max Niemeyer, 1913, p.78.

de schema behavioristă “stimul-reacție”, care ar putea primi și o reprezentare grafică de genul celei de mai jos.



Săgeata din stânga reprezintă intrările în cutia neagră, iar cea din dreapta ieșirile. S-ar părea, la prima vedere, că nu știm – din punct de vedere behaviorist – nimic despre ceea ce se petrece în cutia neagră. Unii spun chiar că, din punctul de vedere al acestei scheme, “mintea” nici nu există. Mie mi se pare mai potrivit să spunem că mintea este pusă între paranteze.

Din punct de vedere filosofic, este extrem de interesant să detectăm presupuzițiile schemei.

Stimulările sunt evenimente la care, conform acestei scheme mintea răspunde cu o reacție. Mai corect spus, asociază un comportament (o secvență de acțiuni ale subiectului uman) cu o intrare. Mecanismul acestor asocieri în minte ar putea să ne și scape. Important din punct de vedere behaviorist este că aceste asocieri se formează. Altfel, de fapt, behavioriștii nu ar avea ce să studieze.

Semnificația acestei observații cred că este mai limpede dacă vom compara mintea behavioristă cu o minte carteziană. Din punct de vedere cartezian, schema de mai sus prezintă o mare problemă, deoarece reacțiile sunt și ele evenimente (din lumea fizică!). Cu alte cuvinte, schema behavioristă presupune că mintea este prinsă în înălțuirile de evenimente din lume. Mai mult decât atât, ea este o parte a lumii fizice; nu este altceva decât un mod de a asocia între ele evenimente. Chiar dacă asocierea respectivă ar fi un accident din punctul de vedere al legilor fizicii, ea tot face parte din ea.

Din perspectivă carteziană, mintea este înăscută. Este înăscută nu neapărat în sensul biologic, ci în sensul că procesele din minte fac parte din altă lume. Ele nu ar putea fi detectate prin observarea mecanismului asocierilor. Desigur, ne putem și naște cu capacitatea de a face anumite asocieri și atunci, din nou, doar vârful aisbergului este vizibil cu ajutorul

schemei behavioriste. Într-un fel sau altul, procesele din minte sunt mai bogate decât presupune schema behavioristă.

Un corolar al presupuziției asociaționiste subiacente schemei behavioriste este și posibilitatea dezvoltării unor tehnologii de modificare a comportamentului. Evenimente selectate convenabil la intrare (de pildă prin adăugarea unor stimulente aidoma celor folosite în dresarea animalelor) ar putea permite dezvoltarea unor tehnici de antrenare a comportamentelor.

Înainte de a discuta însă ideile unui psiholog behaviorist celebru, ar trebui să remarcăm schimbarea intervenită în nivelul analizei. Atunci când am cercetat ideile lui Ryle sau Wittgenstein analiza se mișca în planul argumentelor logice și lingvistice. Acum discuția noastră privește presupuzițiile metodologice empirice behavioriste.

Psihologul la care făceam aluzie mai sus este B. F. Skinner. El a dezvoltat, între altele, și ideea unei tehnologii a comportamentului.<sup>56</sup> Pe ce bază ar urma să fie totuși concepută o astfel de tehnologie? Care este mecanismul schemei de mai sus, după Skinner?

Cheia concepției lui Skinner este că evenimentele din mediu au un rol în selecția comportamentelor:

*...the environment acts in an inconspicuous way: it does not push or pull, it selects.*<sup>57</sup>

Evident, ideea rolului selecției naturale în evoluție provine din biologie. În acest caz, ea este folosită pentru a înțelege evoluția comportamentului. Se coroborează astfel și observația noastră de mai sus că investigația behavioristă se îndreaptă către formarea de comportamente (mai precis, către selectarea asocierilor dintre evenimente din mediu și reacții comportamentale).

---

<sup>56</sup> Cf. B. F. Skinner, *Beyond Freedom and Dignity*, New York, Bantam Books, 1972.

<sup>57</sup> Skinner, *Op.cit.*, p.14.

Divorțul psihologiei de filosofie este un imperativ pentru Skinner. El remarcă faptul că fizica grecilor vechi are doar un interes istoric astăzi. Deplânge situațiile în care scrierile lui Platon sau Aristotel sunt și astăzi recomandate studenților interesați de cercetarea comportamentului. După Skinner, ceea ce spun Platon sau Aristotel despre comportamentul uman are doar un interes istoric.<sup>58</sup>

Defectul explicațiilor mentale, arată Skinner, ține de faptul că ele pun capăt curiozității noastre. "De ce ai fost la teatru?" Dacă am accepta o explicație în termenii psihologiei populare, una de genul pe care ar oferi-o persoana respectivă, "Pentru că am simțit nevoia de a merge la teatru", am pune punct cercetării. Strategia behavioristă ar consta în căutarea în mediu a ceea ce în trecut a determinat mersul persoanei respective la teatru. Ea ar fi mult mai bogată.<sup>59</sup>

Chiar și comportamentul verbal, recunoscut pentru complexitatea sa, urmează să fie studiat după aceeași metodologie. Skinner scrie că noi acordăm credit pentru diversele comportamente verbale: un credit mic dacă repetăm o propoziție pe care am auzit-o cu câteva clipe înainte, un credit mai mare dacă repetăm o propoziție citită acum un an, un credit mult mai mare dacă formulăm o propoziție originală.<sup>60</sup>

Programul filosofico-științific al lui Skinner este vast. El include o concepție despre proiectarea culturilor, discuții despre pedeapsă, libertate și demnitate. Toate acestea depășesc evident obiectivele noastre în această carte și țin de filosofia științelor sociale sau de filosofia politică. Din această perspectivă mai amplă, am putea spune că geniul cel rău imaginat de Descartes și-ar putea atinge scopurile.

Nu este momentul ca să evaluăm metodologia lui Skinner. Pentru a întreține însă interogațiile filosofice, vom menționa totuși o distincție făcută de Daniel Dennett. Dennett arată că ființele skinneriene sunt o extensie a creaturilor

---

<sup>58</sup> Skinner, *Op. cit.*, p.3.

<sup>59</sup> Skinner, *Op. cit.*, p.10.

<sup>60</sup> Skinner, *Op. cit.*, p.45.

darwiniene.<sup>61</sup> Problema este însă că aceste ființe skinneriene se disting de ființele popperiene<sup>62</sup>, care selectează idei.<sup>63</sup> Ființele popperiene pun ideile să moară în locul lor. Riscul pentru ființele skinneriene este de a fi ucise de una dintre erorile lor.

## Filosofie și psihologie

În rezumat, la cumpăna dintre veacurile al XIX-lea și al XX-lea, filosofia și psihologia s-au îndreptat către direcții diferite.

În planul metodelor, al uneltelor folosite de filosofie trebuie menționată cu precădere cultivarea analizei logice a limbajului de către filosofia analitică.

Psihologia, în schimb, a câștigat foarte mult din cultivarea metodelor experimentale.

Repererele de bază ale acestui parcurs divergent sunt date prin urmare de:

- **Analiza logică a limbajului**

*{Analiza logică a devenit o analiză făcută cu mijloacele logicii matematice. Forța acestor unelte i-a făcut, de altfel, pe unii filosofi analitici să creadă că vor putea unifica științele empirice, reducând, între altele, psihologia la fizică.}*

- **Metodologia behavioristă**

*{Behavioriștii cercetează empiric felul în care sunt selectate diversele asocieri dintre stimuli și reacții comportamentale.}*

- **Respingerea explicațiilor mentale de către behavioriști**

<sup>61</sup> Daniel Dennett, *Kinds of Minds*, Londra, Phoenix, 1996, p. 112 (cartea este tradusă și în limba română la Humanitas).

<sup>62</sup> De la numele filosofului Karl R. Popper.

<sup>63</sup> Dennett, *Op.cit.*, p. 116.

*[Explicațiile cu ajutorul entităților mentale sunt respinse pentru că sărăcesc cercetarea. Explicațiile mentale ridică un obstacol în calea investigării lanțului de stimulări și reacții.]*



## Uneltele cele noi ale împăratului?

Din când în când, filosofii, și nu numai ei, speră că au descoperit unelte care îi vor ajuta să rezolve orice problemă. Așa s-a întâmplat în secolul al XX-lea cu logica matematică. Ea a fost adoptată de către mulți filosofi ca un instrument menit să schimbe situația din filosofie. Cu ajutorul noilor metode logice urmau să fie curățate grajdurile augiene ale filosofiei.

Până la un punct operațiunea a și reușit. Mai mult decât atât, noile instrumente au dus la apariția unor noi subiecte de dezbatere filosofică și a unor argumente, formulate nu întotdeauna în limbaj matematic, dar de neconceput înainte de dezvoltarea logicii simbolice. Am să dau în acest sens un exemplu, nu tocmai simplu, dar relevant pentru filosofia minții. Este vorba de un argument formulat de Saul Kripke.

Logica modală, în care și-a făcut inițial un nume Saul Kripke, exista încă din Antichitate. O dată cu apariția noilor unelte însă, filosofii i-au dat și ei o haină matematică. Au extins astfel teritoriile logicii filosofice. Sistemele create inițial aveau însă doar o structură sintactică. Marea descoperire a lui Kripke a constat în găsirea căilor de construire a unor semantici. Filosofii au redescoperit totodată noile valențe ale conceptului de lume posibilă sau rosturile unor termeni precum cel de designator flexibil și designator rigid.

De ce este logica modală relevantă pentru filosofia minții? Pentru a înțelege acest lucru gândiți-vă la teoria conform căreia proprietățile mentale sunt, de fapt, proprietăți ale creierului. În limbajul logicii modale se poate spune că proprietățile mentale sunt identice cu proprietăți ale creierului.<sup>64</sup> Este aceasta o identitate necesară sau una contingentă? Pentru

---

<sup>64</sup> Vezi analiza lui Saul Kripke din "Identity and Necessity", în Milton Munitz (ed.), *Identity and Individuation*, New York, 1971 (fragmente în limba română puteți găsi în *Filosofia analitică*).

cei care propun teoria respectivă identitatea în cauză este contingentă. Este o descoperire științifică faptul că proprietățile mentale sunt proprietăți ale creierului.

Să vedem strategia lui Kripke menită să respingă teoria respectivă. Mai întâi el precizează că un designator rigid este o expresie care desemnează în toate lumile posibile același obiect. De pildă, expresia "rădăcina pătrată a lui 4" se referă în orice lume posibilă același obiect. Apoi se precizează că un enunț necesar este un enunț adevărat în orice lume posibilă (în orice situație posibilă). Un enunț este contingent dacă există situații posibile în care nu este adevărat.

Subtilitatea lui Kripke constă în distingerea planurilor analizei: pe de o parte, distincția necesar-contingent ține de metafizică; pe de alta, distincția, de care vrea să uzeze în continuare, a priori-a posteriori, ține de teoria cunoașterii. Ceva poate să fie necesar și să fie cunoscut a priori,<sup>65</sup> dar poate fi și necesar și să fie cunoscut a posteriori. În acest din urmă caz, exemplul lui Kripke ar fi de genul "aceste taste nu sunt de gheață". O dată ce știu acest lucru nici nu mai pot să-mi închipui cum ar fi putut fi din altceva decât din plastic (nu ar mai fi aceste taste!). Lucrul acesta îl știu însă a posteriori.

Să revenim acum la identități contingente. De ce ar fi identitatea dintre căldură și mișcarea moleculară contingentă? La prima vedere, pentru că există ceva ce se numește "senzație de căldură" și ne putem imagina o lume posibilă în care avem o senzație de căldură, dar această senzație este provocată de altceva decât mișcarea moleculelor. Greșeala aici, crede Kripke, constă în identificarea căldurii cu senzația de căldură. S-ar putea ca să existe ființe de pe alte planete care au senzația noastră de căldură când pun mâna pe gheață. Kripke crede că, de fapt, "căldură" și "mișcare moleculară" sunt designatori rigizi și

---

<sup>65</sup> A priori înseamnă, după Kripke, ceva care poate să fie cunoscut independent de orice experiență (v. Saul Kripke, "Naming and Necessity", în G. Harman și D. Davidson (editori), *Semantics of Natural Languages*, Dordrecht, Reidel, 1972, p. 260.

identitatea dintre căldură și mișcarea moleculară este necesară. Desigur, noi nu știm însă a priori acest lucru.

Ce se întâmplă în filosofia minții? Aici, spre deosebire de raportul dintre senzația de căldură și căldură, senzația de durere trebuie să fie această senzație. Nu pot avea o durere fără senzația de durere. Acum este momentul crucial al argumentului lui Kripke. Expresia care desemnează proprietatea cerebrală corespunzătoare și expresia care se referă la durere sunt ambele designatori rigizi. Enunțul de identitate relevant în acest caz este unul necesar și nu unul contingent, cum ar vrea teoria identității dintre proprietățile cerebrale și proprietățile mentale. Partea proastă pentru teoria identității dintre cerebral și mental este că aici nu s-a recurs la o proprietate contingentă pentru a constata durerea (ca în cazul căldurii).<sup>66</sup>

Aceasta este critica teoriei identității dintre fizic și psihic. Kripke nu pretinde că ar avea și o soluție la problema corp-minte. Și să spun drept nici nu cred că ar putea să o aibă, lucrând numai cu metodele analizei logice. Rezultatul obținut mai sus este interesant în măsura în care ținem cont de faptul că teoria identității între proprietăți mentale și proprietăți cerebrale trebuie să lupte cu obiecția că oamenii știu o mulțime de lucruri despre dureri și alte senzații, dar nu știau sau cei mai mulți nu știu nici azi mare lucru despre creier. Singura cale de ieșire este să arăți că descoperirea identității în cauză este o descoperire științifică. Or, ceea ce a arătat Kripke este tocmai caracterul necesar al identității respective și faptul că durerea și alte stări sunt sesizate cu ajutorul unor proprietăți necesare. Dacă identitatea ar exista, am putea cunoaște funcționarea creierului independent de orice experiență. Evident lucrurile nu stau însă așa. Nu ne așezăm în fotoliu și încercăm să descoperim cum funcționează creierul nostru.

---

<sup>66</sup> Pentru obiecții la demersul lui Kripke vezi trimerile relevante din capitolul "În jungla bibliografiei". De asemenea, a se vedea considerațiile lui David Chalmers în *The Conscious Mind*. Despre Chalmers vezi capitolul "Misterele minții".

Pe de altă parte, limita analizei logice ca atare iese foarte bine în evidență din acest exemplu. Ea nu constă în obiecțiile care s-ar putea aduce argumentelor lui Kripke. Este limpede că se pot aduce contraargumente la ideile cu privire la caracterul necesar la identităților. Am putea căuta tot felul de defecte în argumentele de mai sus. Dar un lucru pare sigur că nu-l putem face: să reconstituim logic funcționarea minții și să și testăm rezultatul reconstrucției noastre. Aceasta este taxa care trebuie plătită în mod necesar pentru divorțul dintre psihologie și filosofie.

Mai există o problemă delicată, a cărei importanță va ieși la iveală doar treptat. Orice reconstrucție logică a mecanismelor minții are o presupuziție insidioasă: centralitatea logicii și a inferențelor în cadrul mecanismelor minții. Cum această presupuziție s-ar putea să fie falsă, vedem și mai limpede prăpastia care se sapă între filosofia minții făcută numai cu uneltele analizei logice și realitate.

## Filosofie și lingvistică

Studierea metodelor logicii moderne nu constituie un obiectiv al acestei cărți. Familiarizarea, cel puțin la un nivel elementar cu ele, a cititoarei sau cititorului ar fi însă de dorit.<sup>67</sup> Este greu să înțelegi evoluțiile din filosofia contemporană fără perspectiva pe care o dă cunoașterea modului în care logica a influențat filosofia.

Consecința de bază a cotiturii în filosofie reprezentată de recursul la metodele logicii moderne a fost interesul pentru

---

<sup>67</sup> Pentru cei care vor o familiarizare rapidă cu terminologia și ideile logicii moderne o bună enciclopedie de filosofie este sursa optimă. Enciclopedia de referință este aici *Routledge Encyclopedia of Philosophy*, Version 1.0, 1998 (versiunea electronică a enciclopediei de filosofie editată de Routledge, accesibilă pe rețeaua locală a Facultății de Filosofie a Universității din București). Versiunea electronică este recomandabilă pentru ușurința cu care poți verifica pe loc, în bibliotecă, dacă ai înțeles un termen sau altul. Se poate începe cu articolul "Logic in the early 20<sup>th</sup> century" de Gregory H. Moore, paragraful "What is modern logic?" (v., de asemenea, "Glossary of logical and mathematical terms").

limbaj manifestat în filosofie. Cred că nu riscăm prea mult dacă spunem că antipsihologismul filosofic a generat un interes pentru limbaj care a trecut cu mult dincolo de cercurile filosofice care utilizează profesional logica modernă.

Acest interes pentru limbaj nu avea cum să nu ducă, mai devreme sau mai târziu, la apropierea dintre lingvistică și filosofie. Se poate spune chiar mai mult decât atât: metodele logicii au fost preluate în lingvistică și au dus la o revoluție în studiul limbilor naturale. Această revoluție este asociată cu numele lui Noam Chomsky.<sup>68</sup> Ea a început în 1957, când Chomsky publica *Syntactic Structures*.

Metodele lingvisticii lui Chomsky sunt departe de a fi triviale, dar partea lor elementară creă că poate înțelege pornind de la exemple nepretențioase. Pentru aceste exemple avem nevoie de o idee foarte generală de regulă.

Dintr-o perspectivă foarte generală, o regulă are “o parte stângă” și “o parte dreaptă”. Intuitiv vorbind, regula ne arată că putem “transforma” partea stângă a regulii în partea dreaptă. Cel mai comod este să folosim o săgeată și să punem “partea stângă” la originea săgeții și partea dreaptă în vârful săgeții.

Să zicem acum că vrem să explicăm unui copil ortograme. Cum scriem “s-a” ori “sa”? Îl putem învăța să facă teste critice prin care înlătură pe “s” și vede care sunt consecințele. Repetă testele până se convinge că soluția este plauzibilă. Suprimarea unui “s” poate fi privită elegant ca o instanțiere (o aplicare a unei reguli). De exemplu:

*Ion s-a dus la munte* → *Ion a dus la munte* \_

Partea dreaptă are sens. Lacuna “\_” ar putea fi completată (urmând altă regulă) prin ceva de genul “o carte”,

<sup>68</sup> Chomsky a scris mult și concepțiile sale lingvistice s-au schimbat de-a lungul timpului. De aceea, este greu să te referi în bloc la o operă atât de complexă. Pentru cititorul român cea mai bună trimitere este la Noam Chomsky, *Cunoașterea limbii*, București, Editura Științifică, 1996, 368 p. Pentru reperele operei lui Chomsky vezi postfața Alexandrei Cornilescu (pp. 315-316).

dar și printr-o referire chiar la Ion. Trebuie deci scris “s-a”. Faptul că partea dreaptă are sens este crucial pentru interpretarea testului. Vezi, prin contrast:

*Cartea sa este ruptă → Cartea a este ruptă*

Partea dreaptă ar fi avut sens dacă l-am fi suprimat pe “sa” cu totul; așa nu are sens. Deci scriem “sa”.

Testele de genul acesta sunt precum experimentele. Trebuie folosite critic. Chiar dacă un test este trecut, s-ar putea ca altul să nu fie trecut. Regulile însă sunt univoce. Ele ne spun cum să transformăm ceva în altceva.<sup>69</sup>

Un set de reguli poate să transforme întotdeauna un enunț corect gramatical în alt enunț gramatical, dar să nu păstreze proprietatea de a avea sens. “Pătratele sunt rotunde”, de exemplu, este un enunț gramatical corect și poate fi obținut folosind un asemenea set de reguli. Spunem că regulile respective reprezintă o sintaxă a limbii, nu o semantică.

Nu este totdeauna ușor să descoperi setul de reguli potrivit. “Pătratele sunt scunde” este un enunț poetic, fără noimă sau trădează necunoașterea categoriei obiectelor scunde? Chomsky s-a ciocnit de acest gen de dificultăți. Dintr-o perspectivă foarte generală, ceea ce a reușit el însă să impună în lingvistică este ideea că a cunoaște o limbă înseamnă a stăpâni un set de reguli.

Setul de reguli care definesc o limbă este finit, dar extrem de complicat. Mintea noastră are însă capacitatea de a stăpâni regulile limbii. Tocmai această capacitate l-a fascinat pe Chomsky. El și-a pus și întrebarea cum ar fi posibil ca un copil să învețe toate aceste reguli prin mecanismele de selectare a comportamentelor verbale, ca în teoria lui Skinner. Răspunsul

<sup>69</sup> În gramatica lui Chomsky, ar trebui precizat, transformările nu se aplică propozițiilor ca atare, ci structurilor de profunzime. Aceste aspecte specifice sunt însă mai puțin importante aici. Ceea ce vrem să surprindem este utilizarea ideii de regulă ca atare.

său a fost răspicat negativ.<sup>70</sup> Chomsky a adoptat o poziție carteziană în filosofia minții și a susținut că stăpânim regulile respective pentru că avem deja în minte, la naștere, o gramatică universală. Ea nu trebuie decât să fie activată și parametrizată, ca să spunem așa, în funcție de mediul lingvistic în care crește copilul.

Din relatarea noastră se vede cum valul stârmit în lingvistică a lovit și în psihologia behavioristă. Treptat aceasta a pierdut pozițiile pe care le deținea și a început să fie înlocuită cu o psihologie zisă cognitivistă, o psihologie care – așa cum o arată și numele – aduce în centrul atenției cunoașterea de care dispune mintea umană. Cunoașterea aceasta poate să fie, de pildă, cunoașterea limbii, stăpânirea regulilor limbii. Mintea, pusă între paranteze de behavioriști, a ajuns din nou în centrul atenției.

Aceasta este deocamdată schița doar a unei povești incomplete. Cred că este puțin probabil ca filosofia centrată pe logică, noua lingvistică și psihologia cognitivistă să fi putut face față inerentului val de obiecții dacă nu ar fi existat un nou set de metode și modele, cele dezvoltate în știința computerelor. Combinarea interdisciplinară a acestor abordări în cadrul a ceea ce poartă numele de “știința cogniției” avea să dea naștere unor noi și extrem de intense speranțe că misterele minții vor putea fi dezlegate. Pentru a înțelege știința cogniției trebuie să avem însă o idee mai exactă despre mașinile de calcul.

## Mașini și reguli

Ce anume a caracterizat cu precădere al doilea mileniu după Hristos? Multe răspunsuri par plauzibile. “Mașinile” – cred eu – reprezintă o ipoteză căreia ar trebui să-i acordăm mai mult credit decât suntem tentați să o facem la prima vedere.

---

<sup>70</sup> N. Chomsky, ‘Review of Verbal Behavior by B.F. Skinner’, *Language*, vol. 35, 1959, pp. 26-58.

Unul dintre motivele puternice pentru care mașinile nu par a fi un candidat de prima mână în topul de mai sus este acela că tehnologia, arta de face mașini, este îndeobște văzută ca o extensie a științei.<sup>71</sup> Dacă ar fi așa, atunci știința ar rămâne poate fără rival în cadrul celui de al doilea mileniu.

O teză crucială pentru argumentul acestui capitol este însă că tehnologia, deși utilizează rezultatele științei, nu este o extensie a acesteia. Pe de o parte, adesea dezvoltarea aparatelor (cele de măsură, de exemplu) este o condiție preliminară a dezvoltării științei experimentale și nu invers. Pe de alta, nu se poate vorbi de “deducerea” planurilor de construcție a mașinilor din teorii științifice. Iluzia că lucrurile ar sta așa este foarte răspândită, dar popularitatea ei nu este și o dovadă ce ar trebui luată în serios.

Ce face însă o mașină? Dacă este să folosim modul de gândire matematic pentru a da un răspuns la această întrebare, atunci ar trebui să reducem problema la altă problemă, până dăm de una al cărei răspuns să-l știm. Dacă prin mașină înțelegem vreunul dintre dispozitivele automatizate pe care le întâlnim în industrie sau în casă, atunci cel mai simplu lucru ar fi să spunem că o mașină face ceea ce ar face o unealtă acționată de mâna omului. Astfel, ideea este să reducem toată întrebarea la o întrebare despre unelte.

Uneltele, din câte ne permit să apreciem dovezile arheologice, le sunt cunoscute oamenilor de peste două milioane și jumătate de ani.<sup>72</sup> Este vorba despre unelte din piatră, folosite la tăiere, cioplire, curățare etc. Mergând pe linia generalizării ideii de unealtă, lor le putem adăuga focul, cunoscut oamenilor de cel puțin un milion și jumătate de ani.<sup>73</sup>

---

<sup>71</sup> Tehnologia, sună definiția de dicționar, este “aplicarea cunoașterii științifice în vederea atingerii scopurilor practice ale vieții umane sau, cum se spune uneori, la schimbarea și manipularea mediului” (v. *Encyclopedia Britannica* CD 1997, “technology”).

<sup>72</sup> V. *Encyclopedia Britannica* CD 1997, art. “Tools”.

<sup>73</sup> Cf. George Basalla, *The Evolution of Technology*, Cambridge University Press, 1988, p.7.



La ce conduce folosirea uneltelor? Privind lucrurile din punctul de vedere cel mai general, observăm că uneltele servesc la desfășurarea unui proces prin care ceva (inputurile, elementele inițiale ale procesului, intrările) este transformat în altceva (outputuri, rezultate finale ale procesului). Focul este, din punctul acesta de vedere, paradigmatic. Dar orice unealtă servește, în fond, aceluiași scop. Iar mașinile complet automatizate nu sunt altceva decât unelte care servesc la declanșarea unor procese în care omul furnizează doar inputurile și beneficiază de outputuri.

Legătura cu regulile este frapantă. Ea are nevoie însă de o serie de clarificări. Pe de o parte, este evident că aplicarea unei reguli declanșează un proces. Problema este totuși cum pot fi studiate regulile în acest sens extrem de general. Dacă ne aplecăm asupra proceselor declanșate de aplicarea de reguli, ne punem imediat întrebarea dacă ele sunt cumva parte a mersului "legic" al lucrurilor. Mai exact spus, le deduce omul din studiul legilor naturii? Sunt ele o parte a efortului de a aplica aceste legi?

Să ne gândim puțin la foc ca unealtă! Dacă pădurea ia foc în urma unui fulger, acesta este un proces natural. Dacă omul sapă o groapă și face focul în ea pentru a frige carne, acesta nu mai este un proces care să fie declanșat spontan în natură. Ceea ce a observat însă omul în cazul fulgerului a fost un accident intervenit în natură. Omul a căutat însă să reproducă în mod sistematic asemenea accidente. S-ar părea deci că lucrurile stau exact invers: nu observăm regularități și încercăm să le reproducem, ci observăm accidente și încercăm să le dăm caracterul unor regularități dirijate (cu ajutorul uneltelor). În ciuda faptului că oamenii folosesc calculele matematice și legile fizicii pentru a proiecta cu precizie uneltele lor, acestea nu sunt deduse din aceste legi.

Nici nu ar fi nimic de dedus. Cana cu care beau apă este făcută din lut ars și are o toartă. Nu are sens să căutăm vreo lege a naturii care să ne arate de ce cana are toartă. Toarta este o soluție la o problemă de ordin tehnic, iar soluția respectivă este bună sau rea din punct de vedere tehnic și nu pentru că "ar

corespunde” legilor naturii sau le-ar folosi mai bine și așa mai departe.

## Mașina universală

Dacă mai sus am făcut o legătură între mașini și reguli, ne putem pune acum întrebarea dacă nu putem stabili vreo conexiune între funcții, mașini și reguli. Conceptul de funcție este un concept matematic. De data aceasta ar fi vorba deci de a explora din punct de vedere teoretic universul regulilor. Inclusiv mașinile despre care va fi vorba vor fi entități teoretice.

Cu ideea de funcție ne vom mai întâlni pe parcurs și nu ar fi rău să-i dăm și ei o caracterizare cât mai generală. O funcție are următoarele componente:

*nume (argumente: tip) {calcul} [rezultat: tip]*

Alegerea parantezelor mai sus nu este întâmplătoare: parantezele rotunde delimitează inputurile; cele drepte outputurile. Funcția are un “nume” cu ajutorul căruia poate fi “chemată” (cam așa cum utilizăm practic o unealtă pentru a face o anumită operație). Ea se aplică unor argumente care sunt de un anumit tip (de pildă, sunt numere). Esențial pentru funcții este că dacă argumentele sunt aceleași rezultatul este același întotdeauna. Cu alte cuvinte, funcția este un fel de unealtă ideală. Pentru a ajunge la rezultat trebuie făcut un *calcul*, care constă în aplicarea unor reguli.

Să zicem că luăm drept tipuri de argumente șiruri de litere (nu numere). Acestea ar putea fi de genul următor:

sda fghjkl, abc, dea,...

Să zicem acum că problema noastră este să înlocuim litera minusculă “a” cu majuscula “A”. Atunci putem defini o regulă în felul următor:

$XaY \rightarrow XAY$

unde X și Y sunt șiruri (finite) de litere ca mai sus (posibil vide). Astfel în “sda fghjkl” X este “sd”, iar Y este

“fghjkl”; în “abc” X este un șir vid, iar Y este “bc”; în “dea” X este “de”, iar Y este vid. Rezultatul aplicării regulii la aceste șiruri este, evident:

sdA fghjkl, Abc”, deA, ...

O dată ce știm care este tipul argumentelor și cum să calculăm ca mai sus rezultatul, putem spune că am definit o funcție care înlocuiește o apariție a lui “a” cu “A” într-un șir de litere.

Definirea funcțiilor aritmetice este mai complicată, dar structura funcțiilor aritmetice ca atare este aceeași. Iar, cu toată simplitatea exemplului de mai sus, procesul de calcul are la bază aceeași idee.

Patru mari logicieni, Kurt Gödel, Alonzo Church, Emil Leon Post și Alan Turing, au fost interesați în anii '30 ai secolului al XX-lea de conceptul de calculabilitate (sau, cum se mai spune, computabilitate) a funcțiilor.<sup>74</sup> Gödel a fost interesat de definirea recursivă a funcțiilor. Post a dezvoltat conceptul de reguli de producție (din care, indirect, este inspirată și prezentarea făcută mai sus regulilor). Church a formulat lambda calculul (de care vom aminti în contextul discuției despre limbajele de programare).

Turing a fost însă cel care s-a orientat către ideea unei mașini care calculează funcții. Mai mult decât atât, Turing a fost atras de ideea construirii practice de computere. Mergând pe urmele acestei abordări specifice lui Turing, în cele ce urmează ne vom concentra atenția structurii computerelor.

Mașina Turing este o mașină abstractă. Spre deosebire însă de alte mașini ea este universală. Orice funcție computabilă poate fi calculată pe mașina Turing. O dată cu Turing se naște astfel ideea unei mașini universale. Mașinile celui de al doilea mileniu (ceasurile, de pildă) puteau face doar anumite lucruri.

<sup>74</sup> V. despre computabilitate în *Routledge Encyclopedia of Philosophy*, disponibilă, după cum spuneam, în biblioteca Facultății de Filosofie în formă electronică, pe rețeaua locală.

Sfârșitul de mileniu aduce cu sine ideea unei mașini care nu este profilată. Ea are, de altfel, și alte trăsături interesante, pe care le vom prezenta în continuare.

## De ce Turing?

Turing este un simbol. Spre deosebire de alți logicieni care au descoperit cam în același timp ideea de funcție computabilă, Turing s-a orientat limpede către construirea de computere. O dată cu realizarea practică de computere aveau să fie dezvoltate și uneltele care fac obiectul capitolului de față.

Recapitulând drumul parcurs, se observă cum se produce deplasarea de la analiza logică la descoperirea sistemelor de reguli care generează limbile naturale, apoi la studierea minții care stăpânește aceste reguli și, în sfârșit, la construirea de mașini capabile să execute reguli:

- **Analiza logică și apropierea dintre filosofie și lingvistică**

*[Logica matematică a permis nu doar clarificarea unor concepte, ci a deschis calea abordării unor noi teme. Metodele logice au căpătat apoi o formă adaptată cerințelor lingvisticii.]*

- **Cunoașterea limbii**

*[Din perspectiva noii lingvistici mintea are un rol de jucat, doar făcând apel la conceptul de minte putem înțelege cum este posibilă stăpânirea unor sisteme de reguli.]*

- **Regulile**

*[Un concept foarte general de regulă permite reunirea sub aceeași carcasă conceptuală a abordării sistemelor logice, lingvistice, a calculului funcțiilor și mașinilor.]*

Știința cogniției în versiunea filosofilor, deși împrumută numeroase idei din știința computerelor, rămâne adesea tributară orizontului logicii. Noile concepte sunt proiectate pe fondul deja cunoscut al logicii. În cele ce urmează, vom încerca să facem însă operația inversă. Prezentarea va fi mai apropiată de știința computerelor și vom încerca să arătăm cum a fost integrată logica în cadrul acestora.

## Cine calculează?

În studiile sale antebelice Turing folosea termenul de “calculator” cu sensul de “omul care calculează”. Mașina Turing este o abstracție a omului care calculează. Chiar și după atâția ani acest sens nu ar trebui poate pierdut din vedere. Metaforele folosite pentru a descrie elemente ale structurii computerelor (“memoria”, de exemplu) trădează, de altfel, originea acestui concept.

Ne vom concentra aici atenția asupra schemei de principiu a unui computer văzut ca mașină universală, renunțând la o prezentare a mașinii Turing. Astfel, în locul contemplării unui concept logic, cititoarea sau cititorul sunt invitați să-și facă o idee despre mașini reale, depășind atitudinea distantă pe care mulți filosofi par să o aibă față de aceste mașini. Ideea noastră este că filosofii sunt precum Mary din experimentul lui Frank Jackson: trăiesc în lumea conceptelor pur logice, dar au ceva nou de aflat dacă ies și văd culoarea obiectelor reale.

Prin convenție, vom numi *dispozitive* mașinile care intră în componența altor mașini. Vom descrie mașinile sau dispozitivele folosind o notația următoare:

Numele dispozitivului

< componente >

(intrări)

{ operații }

[ ieșiri ]

În principiu, putem folosi săgeți pentru a arăta felul în care se desfășoară fluxul proceselor în mașină. De exemplu

dispozitiv1 → dispozitiv2

indică faptul că ieșirile din dispozitiv1 devin intrări pentru dispozitiv2. Diagramele cu săgeți au, ce-i drept, avantajul de a vizualiza fluxul proceselor, dar sunt periculoase pentru că

incurajează conceperea unor mașini tip spaghete, în care se sare de la un dispozitiv la altul în mod necondiționat.

## Rastere

Luați un ziar cu ilustrații alb-negru și studiați imaginile cu ajutorul unei lupe suficient de puternice. Veți vedea o mulțime de puncte, mai dese sau mai rare. Până la urmă, ajungeți la concluzia că pozele din publicație sunt în fond făcute din puncte albe și negre. Dacă aveți un ziar mai vechi, puteți chiar vedea cu ochiul liber astfel de puncte. Imaginea ar putea fi astfel descrisă în limbaj cartezian foarte simplu: fiecare punct de dimensiune minimă are proprietatea de a fi alb sau negru; dacă este alb punem un "0"; dacă este negru, punem "1". Astfel, tripletele de numere (primele două pentru coordonatele carteziane și ultimul pentru proprietatea de a fi alb sau negru) pentru fiecare punct ne furnizează datele de care avem nevoie pentru a descrie imaginea.

Dar dacă imaginea este în culori? Există și în acest caz o strategie de a reduce totul la "0" și "1". Facem, de pildă, o listă a culorilor și punem "1" dacă aceasta este culoarea punctului.

Mai general, putem face o listă cu proprietăți pentru orice obiect pe care vrem să-l reprezentăm; punem "1" dacă este adevărat că obiectul are proprietatea respectivă sau "0" dacă obiectul nu are proprietatea respectivă. Cum aceste semne, cifrele binare "0" și "1", se numesc uzual biți, putem spune că proprietățile sunt reprezentate cu ajutorul unei liste de biți.

## Dispozitive simple de prelucrat liste de biți și memorii

Cel mai simplu dispozitiv ar fi cel care transmite mai departe lista de biți așa cum este, fără să schimbe nimic. Este ceva, la urma urmei, ca un cablu sau o conexiune capabilă să

păstreze biții așa cum sunt. Fără îndoială este vorba despre un dispozitiv folositor, dar nu deosebit de interesant.

Un dispozitiv foarte simplu pe care ni-l putem imagina este inversorul. Îl putem descrie în felul următor:

### **Inversor**

(intrări: liste de biți)

{fiecare "0" este transformat în "1" și fiecare "1" în "0"}

[rezultatul este tot o listă de biți]

Lipsa componentelor indică faptul că este vorba de un dispozitiv simplu. Evident, nu suntem obligați să construim dispozitivele reale în acest fel.

Deși simplu, acest dispozitiv poate face lucruri interesante. Închipuiți-vă că lista de biți nu este altceva decât rasterul unei imagini alb-negru. Dispozitivul acesta ne permite să construim imaginea inversă.

Să zicem însă că vrem să comparăm două imagini. Atunci ne putem sluji de un

### **Comparator**

(preia la intrare două liste de biți)

{dacă listele sunt inegale, rezultatul este automat 0; dacă listele sunt egale, vom construi o listă nouă și pe fiecare poziție vom pune 0, dacă biții din listele de la intrare sunt diferiți, și 1, dacă biții sunt identici}

[rezultatul este 1, dacă pe lista nouă nu se găsește decât cifra binară 1; altfel este 0]

Dacă ați învățat logică, v-ați întrebat poate la ce este bun "sau exclusiv" (simbolic XOR). Calculați XOR pentru fiecare poziție a două liste de genul celor de mai sus. Dacă biții sunt diferiți, obțineți 1. Este suficient să folosiți negația (dispozitivul corespunzător nu este altceva decât un inversor) și veți realiza

un comparator.<sup>75</sup> În continuare vom vedea ce importanță au comparatoarele în construirea unor dispozitive de calcul complexe.

Memoriile sunt, la urma urmelor, tot niște dispozitive, necesare pentru sisteme mai complexe. De ce? Pentru a putea depozita rezultatele unor prelucrări. Posibilitatea de a depozita rezultatele prelucrărilor este cheia construirii unui sistem extrem de flexibil, în care diferite dispozitive pot intra în acțiune unul după altul.

Cel mai simplu mod de a ne reprezenta o memorie este cel al unui sistem de fișe. Probabil fiecare cititoare sau cititor și-a făcut fișe pentru a nota extrase din lucrările citite, cuvinte, explicații ale acestora și așa mai departe. Secretul memoriei este să poți și să regăsești ce ai notat pe fișe. Un sistem de adrese este absolut necesar. Schematic, dispozitivul ar arăta astfel:

### Memorie

(semnal care indică dacă trebuie introdusă sau extrasă o listă de biți; adresa; dacă este cazul, lista de biți)

{după caz, stocarea, copierea sau ștergerea listei de biți de la o anumită adresă}

[confirmarea succesului operației de stocare, copiere sau ștergere; dacă este cazul, lista de biți copiată]

Cheia unei astfel de memorii pentru liste de biți este dată de adresele care sunt și ele tot liste de biți. Într-o memorie pot fi stocate deci și adrese. Evident, aceste adrese pot fi și ele prelucrate.

---

<sup>75</sup> Pentru o descriere mai tehnică a dispozitivelor care implementează funcții booleene v. Andrew S. Tannenbaum, *Organizarea structurată a calculatoarelor*, ed. a IV-a, București, Computer Agora Press, 1999, pp.109-125 (pentru comparator v. pp.123-124).



## Dispozitive care declanșează alte dispozitive

Cel mai simplu mod de a pune în funcțiune dispozitive ar fi reprezentate de montarea lor în serie și acționarea pe măsură ce fluxul de biți ar trece prin dispozitivele respective. Dezavantajul unui asemenea montaj este evident. Pentru a schimba mașinăria ar trebui refăcut (fizic) montajul.

Ideea de bază este de a avea un dispozitiv de genul unui

**Dirijor**

(preia liste cu coduri ale dispozitivelor)

{produce semnale către dispozitive}

[semnalele declanșează acțiunea dispozitivelor]

Dirijorul acționează cam ca un dirijor al unei orchestre: se uită pe partitură și face semn părții din orchestră care trebuie să intre în acțiune.

Dirijorul poate să execute, ca și montajul descris inițial aici, o declanșare în serie a diverselor dispozitive (în ordinea indicată pe lista de intrare în dispozitivul de dirijare). Dirijorul are, pe lângă aceasta, alte posibilități: el poate să sară de la un dispozitiv la altul. De exemplu, un dispozitiv produce un anumit rezultat, iar rezultatul este dirijat către un alt dispozitiv, care nu este neapărat cel aflat fizic în prelungirea primului. Dirijorul stochează rezultatul în memorie și de acolo îl trimite către următorul dispozitiv, făcând un salt peste alte dispozitive.

Fundamental este faptul că salturile pot fi de două tipuri: necondiționate sau condiționate. Dacă saltul este necondiționat, dirijorul reorientează pur și simplu fluxul de biți către alt dispozitiv. Dacă saltul este condiționat, atunci acțiunea dirijorului implică o folosire mai complexă a memoriei și a unui comparator.

Rolul comparatorului este acela de a furniza dirijorului date cu privire la listele de biți supuse prelucrării. Este lista la fel ca altă listă? Este mai lungă, mai scurtă și așa mai departe?

Rezultatele comparatorului sunt stocate la o adresă de memorie și sunt folosite de către dirijor pentru a stabili dacă o anume condiție este satisfăcută (de exemplu, dacă două liste sunt identice) și a orienta fluxul de biți către un dispozitiv sau altul.

Marele avantaj al dispozitivelor de dirijare este reprezentat de obținerea unei mașinării extrem de flexibile, dar care nu este montată și remontată pentru a executa diverse acțiuni.

## Sistemele complexe de prelucrat biți

Prezentarea abstractă de mai sus vă poate duce cu gândul la structura calculatoarelor. Intenția este însă mai generală. Arhitectura calculatoarelor este doar un caz particular. Înainte de a o discuta puțin, este necesar să mai facem câteva precizări.

De ce biți? Simplificarea pare foarte mare. De ce nu sisteme ternare sau mai complicate? Pe de o parte răspunsul ține de considerente practice: sistemele electronice sunt mult mai fiabile dacă folosesc biți (de exemplu, tensiune nulă pentru bitul 0 și o tensiune nenulă pentru bitul 1). Pe de altă parte există, din punct de vedere teoretic, considerente mult mai importante: lumea poate fi reprezentată cu ajutorul unor rastere (care înglobează datele referitoare la proprietăți). Transformarea acestor rastere nu este altceva decât o reprezentare a transformărilor pe care le suferă proprietățile și, implicit, lumea.

Arhitectura calculatoarelor uzuale, la care au contribuit geniul unor cercetători precum Alan Turing sau John von Neumann, este însă extrem de instructivă. Ea ne dă o idee clară despre felul în care arată organizarea structurală a unui sistem complex.<sup>76</sup>

---

<sup>76</sup> Desigur, pentru a adânci subiectul sau pentru o bună documentare, trebuie mers la o lucrare specială. Cartea mea favorită este cea a lui Andrew Tannenbaum, citată mai sus. La această lucrare celebră aș trimite pentru o perspectivă serioasă asupra ideii de arhitectură a unui sistem de prelucrare de date.

Calculatoarele au un dispozitiv complex numit procesor, care – în realitate – poate îngloba milioane de dispozitive simple. Schematic, procesorul arată astfel:

### Procesor

<memorii, dispozitive aritmetico-logice, dispozitive de dirijare>

(intrările sunt liste de biți)

{procesorul poate stoca în memoriile sale interne, de lucru listele care trebuie prelucrate, codurile dispozitivelor logico-aritmetice și de dirijare care trebuie puse în funcțiune și adrese ale memoriei externe; în esență, procesorul acționează în doi timpi: preia codul dispozitivului, preia lista de biți care trebuie prelucrată sau examinată – dacă este vorba de o condiție pentru un dispozitiv de dirijare - și o trimite către dispozitiv}

[rezultatele sunt liste de biți]

Se observă că și codurile dispozitivelor sunt tot de formă binară. Memoriile interne ale procesorului se numesc, de regulă, “registri”. Există astfel un registru pentru codul următorului dispozitiv care intră în funcțiune.

“Partitura” după care cântă toată această orchestră înglobată în procesor este, evident, o lungă listă de biți. Cum deosebește atunci procesorul codurile dispozitivelor de listele care trebuie prelucrate? Existența unor tacte ale procesorului este cheia. Noi am simplificat lucrurile mai sus și am presupus că într-un tact este adus din memoria externă codul dispozitivului și în al doilea tact sunt preluate liste de prelucrat.

Partitura aceasta se numește program și este, la rândul ei, stocată în memoria internă a computerului, cum li se mai spune calculatoarelor.

Schema de principiu a unui computer este atunci:

### Computer

<procesor, memorie, dispozitive pentru intrarea, circularea și ieșirea listelor de biți>

(intră liste de biți)

{listele de biți sunt stocate în memorie și sunt prelucrate cu ajutorul procesorului}

[ies liste de biți]

Așa numitele dispozitive periferice nu fac decât să transforme ceva în biți (de exemplu, apăsarea tastelor de către mine când scriu acum acest text) sau să preschimbe listele de biți în ceva mai accesibil simțurilor noastre (de pildă, o imagine pe un ecran, un text tipărit, o acțiune a unui robot).

Cei care sunt familiarizați cu capitolele din logică dedicate mașinii Turing și funcțiilor recursive știu desigur că un computer ca acela de mai sus este capabil să calculeze, în principiu, orice funcție. Adăugarea mai multor procesoare, activarea simultană a mai multor dispozitive nu ar conduce la calcularea mai multor funcții. Toate acestea sunt rezultate importante din punct de vedere logic.

Viteza cu care calculează un computer poate fi însă un factor important. Faptul că o funcție va fi calculată în câteva secunde sau rezultatul va veni peste o sută de ani nu este lipsit de semnificație când de calculul respectiv depinde prelucrarea unei imagini pe care vreau să o văd imediat ș.a.m.d.

Metodele de a construi programe care dau rezultate palpabile, așa cum ne oferă și mintea umană, constituie – de fapt – panoplia de unelte noi despre care este vorba în acest capitol. A face deosebirea între rezultatele pur logice și arta de a construi sisteme complexe constituie miezul argumentului acestui capitol.

## Arhitectura computerelor

În rezumat, termenul de “arhitectură”, deși adesea utilizat (nu doar în cazul computerelor, ci și al minții) este foarte înșelător. Cel puțin în cazul meu el îmi sugerează asociații cu structura unor construcții care au prea puțină flexibilitate: clădiri, palate, poate mobilierul. Chiar dacă interioarele clădirilor se pot schimba, tot nu simt flexibilitatea totală ca fiind o caracteristică a lor.

Mașinăriile descrise mai sus sunt, prin contrast, extrem de flexibile. Universalitatea le este conferită de această posibilitate de a-și modifica funcționarea fără a schimba montajele palpabile fizic (așa-numitul *hard*). În același timp, materialul pe care îl prelucreează este uluitor de simplu.

- **Listele de biți**

*{Materialul prelucrat este format din șiruri de cifre binare; acestea sunt folosite pentru a reprezenta proprietăți, dar și ca adrese, ca nume (coduri) ale unor dispozitive.}*

- **Dispozitive de prelucrare a listelor de biți**

*{Biții pot fi prelucrați ca și "0" și "1" din logică sau se poate opera cu șirurile de biți precum cu numerele notate în sistem binar.}*

- **Memorii**

*{Rezultatele operațiilor logice sau aritmetice sunt stocate în memorii, de unde pot fi trimise spre prelucrare altor dispozitive}*

- **Dispozitive de dirijare**

*{Flexibilitatea sistemelor complexe este dată de dispozitivele care pot declanșa, condiționat sau necondiționat, în diverse ordini alte dispozitive}*

Computerele obișnuite au o unitate centrală de procesare, care prelucreează biții și dirijează funcționarea dispozitivelor de prelucrare, după un program stocat în memorie.

## **Limbaj-mașină cu față umană**

La urma urmei, care este rostul sistemului complex ale cărui principii de construcție și funcționare au fost descrise mai sus? Pe de o parte, acesta nu este altceva decât o versiune a conceptului uzual de computer (dacă mergem pe ideea procesării centralizate a biților). Pe de altă parte, este scheletul unui sistem capabil să rezolve probleme dintre cele mai variate (dacă are programul adecvat stocat în memorie).

Ce înseamnă oare “a rezolva probleme dintre cele mai variate”. Închipuiți-vă ce dificultăți ar avea motanul fetei mele dacă ar porni să-și procure de unul singur conserve de carne pentru pisici. Ar trebui să identifice magazinul unde se găsesc conservele, să-și procure cumva cutiile, să le deschidă. Toate acestea cer așa-numita “cunoaștere”, care nu este altceva decât capacitatea de a recunoaște locul unde este magazinul, de face rost de cutii, de a le deschide. Chiar dacă nu ar face lucrurile acestea în chip uman, sistemul complex descris mai sus le poate face în principiu (dacă este dotat, evident, și cu dispozitivele periferice necesare). Singura problemă este că ar trebui introdus un program adecvat în memoria sistemului.

Ca să rezolvăm, tot imaginar, chestiunea introducerii programului, amintiți-vă o clipă de felul în care se lucra mai demult cu centralele telefonice. Poate ați văzut așa-ceva prin filme: centraliste care au în față un panou în care intră zeci de fire și care introduc foarte repede în contacte mufe, pentru a realiza legătura cerută. Ceva asemănător ar putea realiza un operator (rețineți termenul de “operator”) care ar modifica memoria sistemului în mod adecvat (ar scrie în memorie, la adresa potrivită) biții care alcătuiesc programul.

Nouă ne-ar fi greu să urmărim ce face operatorul din pricina dificultății de a lucra direct cu biți. Un translator ne-ar putea face însă următorul serviciu. Noi scriem programul într-o notăție inteligibilă pentru noi, pe care translatorul o traduce apoi în liste de biți utilizabile de către operator pentru a pune în funcțiune sistemul.

Notăția folosită în acest caz se numește “limbaj mașină”. Asemenea limbaje se folosesc pentru a programa computerele. Ele reflectă, desigur, structura dispozitivelor reale înglobate în computer. În cazul nostru, este suficient ca limbajul mașină să fie relevant pentru dispozitivele noastre imaginare. Notăția folosită de noi aici este adaptată nevoilor textului de față și recurge la cuvinte românești. Programul de computer folosit însă pentru a ilustra ideea de limbaj mașină este cel din primul

capitol al cărții lui Samuel Kamin, *Programming Languages: An Interpreter-Based Approach*, Addison-Wesley, 1990.<sup>77</sup>

Limbajul va fi introdus cu ajutorul unor exemple simple. Să zicem, pentru început, că vreți să adunați două numere (2 și 7, să zicem). Atunci programul va arăta în felul următor:

(ADUNA 2 7)

Cuvântul “ADUNA” nu ar trebui înțeles ca o “comandă” dată sistemului. Translatorul traduce, de fapt, cuvântul “ADUNA” într-un cod, codul sumatorului, dispozitivul care adună două numere. Operatorul plasează codul în memorie. Parantezele “( )” pot fi interpretate, din punctul de vedere al operatorului, ca niște indicații pentru setarea adecvată a dispozitivelor de dirijare. Din punctul nostru de vedere, cel al utilizatorului, parantezele rotunde indică faptul că este vorba de intrări: codul dispozitivului care trebuie declanșat, urmat de intrările în acest dispozitiv. Pe parcurs, am utilizat, de altfel, paranteze rotunde pentru a indica intrările.

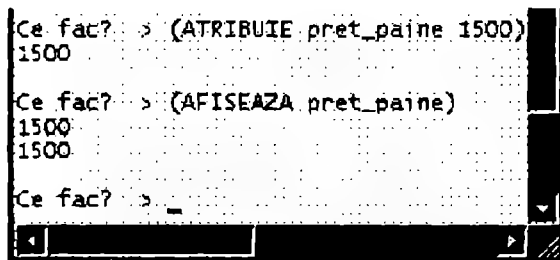
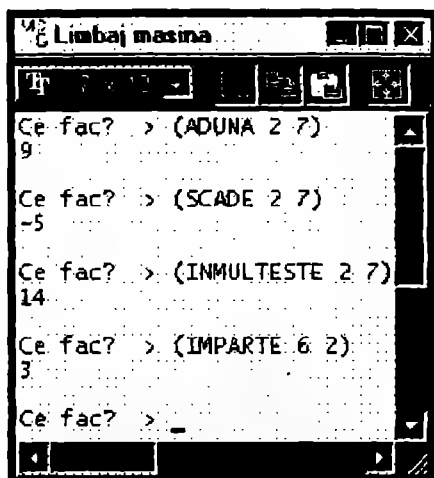
Programul de calculator nu face decât să ilustreze modul de lucru cu un asemenea sistem, după cum se vede în imaginea de mai jos. Programul simulează întrebarea operatorului: “Ce fac?” (ce acțiune întreprind în continuare?) Iar cu operatorul ne închipuim că vorbim prin translator, deoarece noi nu știm ce coduri binare să-i furnizăm acestuia.

---

<sup>77</sup> Programele din cartea lui Kamin se găsesc pe web (sursa distribuției programelor este în directorul [ftp://pub/csci450/95fall/kamin\\_interpreters/distr](ftp://pub/csci450/95fall/kamin_interpreters/distr) de pe site-ul <ftp://ftp.cs.olemiss.edu>. Există și o versiune în Turbo Pascal plasată în directorul [pub/csci450/95fall/kamin\\_interpreters/turbo/](ftp://pub/csci450/95fall/kamin_interpreters/turbo/). Eu am lucrat cu distribuția de bază. Atenție, distribuția de bază conține programe Pascal care nu pot fi compilate fără modificări cu popularele compilatoare Turbo Pascal! O versiune în limbajul C++ a acestor programe ale lui Kamin a fost scrisă de către Tim Budd. Ea este, de asemenea, disponibilă pe Internet.

Rezultatele sunt afișate imediat după ce operatorul a pus în funcțiune dispozitivele de adunat, scăzut, înmulțit și împărțit.

Putem ruga, de asemenea, operatorul să introducă valori (de pildă, numere) la o adresă din memorie. Din nou, noi ignorăm formatul concret al adresei. Nu vom folosi biți, ci denumiri uzuale pe care le putem ține ușor minte. În sfârșit, valorile stocate în memorie pot fi extrase și afișate la nevoie pe ecranul computerului.



Cifra "1500" apare de două ori după ce i-am spus operatorului să afișeze, pentru că sistemul

are un ecou la apeluri. Mai întâi apare ecoul, apoi rezultatul. Marele avantaj al construcției de mai sus este însă că ea ne permite să utilizăm variabile. O variabilă este un locaș (de memorie, în acest caz; sau un loc gol pe o fișă cu care lucrăm, sau o lacună într-un text) care își schimbă pe parcursul procesului de calcul conținutul, ca în imaginea următoare, unde prețul se modifică în urma adăugării unui spor generat de către inflație.



```

Ce fac? > (ATRIBUIE inflatie 500)
500
Ce fac? > (ATRIBUIE pret_nou (ADUNA pret_paine inflatie));
2000

```

Rezultatul nu este amuzant pentru bugetul meu, dar este o ilustrare instructivă a felului în care decurge procesul de calcul, pe care-l vom denumi pe scurt comutație, în sistem. Secvențele din parantezele rotunde sunt ca un fel de atomi de comutație, dacă nu conțin, la rândul lor, paranteze rotunde. Se observă, de asemenea, că o paranteză rotundă poate fi cuibărită în altă paranteză rotundă.

Puterea sistemului este dată de posibilitatea de a defini noi secvențe computaționale și de dirijarea comutațiilor. Cele mai curioase cititoare și cei mai curioși cititori pot încerca, dacă știu limbajul Pascal sau C++, să citească sau, mai bine, să rescrie programul lui Samuel Kamin. Atunci vor înțelege mai bine mecanismul definițiilor. Așa, oarecum din afară, putem preciza că ideea este aceea de a furniza sistemului un cod pentru un dispozitiv nou (care nu există fizic!), o indicație cu privire la intrările în acest dispozitiv și un program care arată felul în care funcționează dispozitivul. Astfel, sistemul poate genera un rezultat ca și cum ar apela la un dispozitiv existent efectiv. De fapt, se face apel la dispozitivele deja cunoscute. Ilustrarea simplă produsă în continuare ne arată felul cum se poate introduce un dispozitiv care să calculeze succesorul unui număr.

```

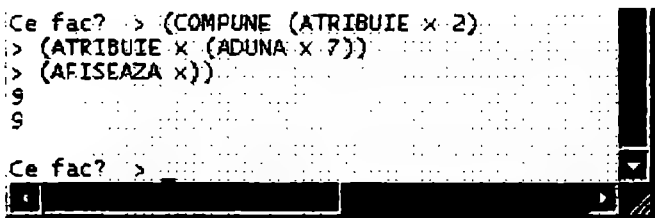
Ce fac? > (DEFINITIE SUCCESOR (x) (ADUNA x 1))
SUCCESOR
Ce fac? > (ATRIBUIE x 0)
0
Ce fac? > (SUCCESOR x)
1
Ce fac? >

```

“DEFINITIE” este un cuvânt-cheie care indică faptul că trebuie creat un cod nou de dispozitiv (cărui translatorul îi asociază cuvântul “SUCCESOR”) și trebuie rezervat un locaș de memorie pentru intrarea aferentă. În sfârșit, micul program “(ADUNA x 1)” trebuie și el stocat undeva în memorie și trebuie apelat atunci când este invocat “SUCCESOR”-ul. Se

observă că acum nu va fi pus în funcțiune un dispozitiv anume, ci va fi rulat un program, care furnizează un rezultat, ca și cum ar exista un dispozitiv special construit pentru a calcula succesorul unui număr.

Telefonistele din filme schimbau mereu legăturile făcute cu tot felul de fire și astfel modificau montajul ca atare. Nu ar fi posibil însă să se schimbe doar funcționarea sistemului? Da. Lucrul acesta îl fac dispozitivele de dirijare. Cel mai simplu lucru este să compunem micile programe care invocă dispozitive.



“COMPUNE” invocă un mod de dirijare secvențial, în care fiecare computație este executată când îi vine rândul. Ce am obținut? Numărul întrebărilor “Ce fac?” ale operatorului s-a redus la una singură. Semnul “>” pus de operator pe ecran indică trecerea pe un nou rând, nu așteptarea de noi instrucțiuni. La nevoie, putem trimite programul scris pe o hârtie. Translatorul îl traduce. Operatorul introduce codurile necesare și utilizatorul primește rezultatul tot pe o hârtie. Utilizatorul nu vede sistemul, nu-l vede pe operator, nu vede translatorul. Circulă doar niște hârtii.

Dirijarea poate să nu fie doar secvențială. Se poate face – cum am mai arătat – o selecție în funcție de condiții date. În cazul condițiilor, se recurge la un comparator care furnizează, cum se spune, un rezultat boolean.<sup>78</sup>

<sup>78</sup> De la numele lui George Boole, inventatorul logicii matematice a propozițiilor (adevărata sau false). Aici vor fi folosite valorile binare, ca în bine-cunoscutele tabele de calcul a valorii logice a propoziției.

```

Ce fac? > (MAIMARE x 10)
0
Ce fac? > (MAIMIC x 10)
1
Ce fac? >

```

Mai poate fi folosită și o computație de genul “(EGAL x y)”. Rezultatele astfel obținute pot fi folosite de un dispozitiv de dirijare care operează selecția în funcție de condiția specificată: dacă este adevărată condiția, atunci dirijează fluxul computației către primul dintre programele între care are de ales; dacă este falsă îl rulează pe al doilea.

```

Ce fac? > (AFISEAZA x)
9
9
Ce fac? > (DACA (MAIMIC x 10)
> (ATRIBUIE x (SUCCESOR x))
> (ATRIBUIE x 0))
10
Ce fac? >

```

Am verificat valoarea lui x în sistem. Am introdus apoi un program care folosește selecția. Se observă că rezultatul a fost afișat cu ajutorul ecoului.

Mai departe? Străduința programatorilor a fost să elimine salturile necondiționate, care ar genera programe încălcite ca niște spaghete. De aici recursul la o dirijare cu ajutorul unor dispozitive care execută cicluri cât timp este adevărată o condiție.<sup>79</sup>

<sup>79</sup> Ada Lovelace este considerată descoperitoarea programării tocmai pentru că, între altele, a introdus utilizarea ciclurilor. Astfel, programele pot fi efectiv construite. Un program care este compus din milioane de alte programe, după schema de mai sus, nu ar putea fi scris pe hârtie fără cicluri. Ce ar însemna să calculăm succesorul de câteva milioane de ori fără a recurge la cicluri?

```

Ce fac? > (ATRIBUIE y 0)
0
Ce fac? > (CATTIMP (MAIMIC y x)
> (COMPUNE (ATRIBUIE y (SUCCESOR y))
> (AFISEAZA y)))
1
2
3
4
5
6
7
8
9
10
0
Ce fac? > (AFISEAZA y)
10
10
Ce fac? >

```

Ne-am folosit, cum se vede, de valoarea preexistentă a lui  $x$  (care era 10). Programul care folosește ciclul "CATTIMP" afișează succesorul lui  $y$  atâta timp cât  $y$  este mai mic decât  $x$ . Ultimul 0 pus mai sus este un ecou al verificării condiției și ne arată de ce s-a ieșit din ciclu.

Un program care nu este în stare să iasă din ciclu blochează sistemul. Sistemul pierde, de fapt, proprietatea sa de a avea o funcționare extrem de flexibilă. El nu va face decât să execute la nesfârșit același lucru. Încercați să puneți drept condiție pe (EGAL 1 1)! Eu unul nu încerc. Nu de alta, dar vreau să-mi termin de scris cartea și nu să aștept să se termine afișarea șirului numerelor naturale.<sup>80</sup>

Asta este tot ce era de spus despre limbajul mașină. Chiar dacă noi am folosit ceea ce am denumit în titlul

<sup>80</sup> Evident, lucrul acesta se întâmplă în aceste sisteme ideale. Într-un sistem real, după cum se poate vedea relativ repede cu un program de genul celui de mai sus, se ajunge la o eroare. Eroarea este dată de limite de memorie. Chiar dacă am curăța memoria de numerele deja obținute și afișate, am obține cu timpul numere prea mari pentru capacitățile memoriilor reale. Sunt și alte probleme, dar ele ar putea fi înțelese doar examinând felul în care este construit programul care simulează sistemul ideal.

paragrafului "limbaj mașină cu față umană", cheia utilizării acestui limbaj stă în plasarea biților potriviți la locul potrivit, astfel încât să fie transformată funcționarea sistemului. Mereu pot fi folosite noi intrări, pot fi imaginate noi dispozitive și procese foarte lungi pot fi pornite cu ajutorul unor semne puse pe o hârtie.

## Sisteme de operare și mașini virtuale

• Până acum ne-am imaginat că operatorul este un om. Desigur, un om cu o memorie aparte și o capacitate de a manipula liste de biți pe care eu unul nu le am. Nu l-am putea înlocui însă cu un program? Ba da! Acest program se numește sistem de operare. Nu este chiar programul de care mă ciocnesc când deschid computerul de la catedră. Computerul are, de fapt, în memorie (din construcție) un program care încarcă sisteme de operare. Așa se face că sunt invitat să aleg sistemul de operare cu care doresc să lucrez.

Ce este un sistem de operare? Andrew Tannenbaum și Albert Woodhull scriu că sistemele de operare pot fi privite din două perspective: ca manageri ai resurselor sistemului sau ca mașini extinse.<sup>81</sup>

Perspectiva cea mai interesantă din punct de vedere filosofic este cea a extinderii sistemului. Am văzut, de altfel, cum – prin programare – pot fi introduse noi dispozitive în sistem (dispozitive care nu există ca montaje fizice distincte în sistem). Practic, nenumărate noi dispozitive pot fi introduse în acest fel. Sistemul de operare face un pas în plus față de introducerea de dispozitive care efectuează operații distincte și relativ simple; el poate să creeze întregi asemenea sisteme de dispozitive virtuale. Pe scurt aceste sisteme se numesc mașini virtuale.

---

<sup>81</sup> Vezi Andrew S. Tannenbaum și Albert S. Woodhull, *Operating Systems*, ediția a II-a, New Jersey, Prentice Hall, 1997, p.44 (vezi tot Capitolul 1 pentru detalii).

Mașinile virtuale extind posibilitățile sistemului de bază. De multe ori scopul este acela de a oferi o față mai prietenoasă utilizatorului. De exemplu, serviciile translatorului nostru imaginar sunt făcute de către programe. Există două mari tipuri de asemenea programe: interpretoare și compilatoare. Un interpretor, practic, face un fel de translație simultană. El traduce programul în limbajul mașinii pe măsură ce "aude" textul (pe porțiuni, cu alte cuvinte). Compilatorul lucrează ca un translator care a primit tot textul și tălmăcește tot programul în limbaj mașină. Rezultatul muncii compilatorului poate fi folosit chiar și atunci când translatorul nu este de față. De prezența interpretorului, evident, nu ne putem dispensa atunci când vrem să rulăm programul.

Conceptul central al unui sistem de operare este cel de **proces**. Un proces este – în linii mari – un program care este executat.<sup>82</sup> Cu alte cuvinte, un proces este ceea ce se întâmplă atunci când sistemul de operare plasează biți la diverse adrese din memorie și începe să declanșeze acțiunea dispozitivelor care prelucrează biții.

În terminologia sistemelor de operare este preferat termenul de "proces". În filosofia minții și știința cogniției ne vom ciocni adesea de termenul de "computație". Am precizat în capitolul introductiv care va fi strategia folosirii acestor termeni.<sup>83</sup> Evident, dacă în cazul unei computații de genul "(ADUNA 2 7)" ne interesează și felul în care sunt folosite adresele din memorie, evenimentele care au loc pe parcurs până la obținerea rezultatului și așa mai departe, atunci atenția noastră este centrată asupra procesului. Nu vom da însă vreo mare greutate teoretică aici distincțiilor dintre acești termeni. "Computație" are mai degrabă avantajul de a nu stârni ecouri metafizice dubioase. "Proces", dacă este utilizat într-un context adecvat, pare însă foarte potrivit și sugerează – într-un fel -- "natura" mașinilor virtuale. Acestea există doar atâta timp cât un anume proces are loc în sistem. Dacă procesul se întrerupe, mașina virtuală dispare și ea.

<sup>82</sup> Andrew S. Tannenbaum și Albert S. Woodhull, *Op. cit.*, p.15.

<sup>83</sup> Practic, vom folosi, de regulă, sintagma "proces computațional".

Eu am început, de pildă, să învăț practic (nu teoretic<sup>84</sup>) programarea pe un computer bazat pe un procesor Z80 – un computer de tip Spectrum, foarte popular la începutul anilor 1980. Computerul acesta diferă fizic extrem de mult de un PC de la sfârșit de mileniu. El poate fi însă recreat ca mașină virtuală pe un PC. Mașina aceasta virtuală dispare de îndată ce nu mai rulez programul care simulează Spectrum-ul. Dacă ar examina cineva PC-ul după ce a fost rulat programul care simulează Spectrum-ul, nu și-ar putea da seama de faptul că am folosit respectiva mașină virtuală, altfel decât prin “bunăvoința” sistemului de operare. Acesta mai păstrează unele date despre programele rulate, uneori ține evidența anumitor caracteristici ale acestora. S-ar putea însă foarte bine ca să nu rămână nici o urmă.

În final, doar o notă despre sistemele de operare reale. Cele mai importante în momentul de față sunt cele de tip UNIX și WINDOWS. Deosebiriile dintre ele sunt irelevante aici. Atâta doar că unele sisteme UNIX sunt foarte utile pentru cei curioși să le vadă structura. MINIX, scris de Tannenbaum, este probabil ideal din acest punct de vedere. De asemenea, poate fi examinată structura mai complexului LINUX. Acest lucru nu este posibil în cazul sistemului WINDOWS. Structura internă a acestui sistem este secret comercial. Noi folosim doar varianta compilată, evident inteligibilă doar pentru computere. Dacă aveți însă acces la documentația WINDOWS puteți citi foarte multe lucruri interesante despre mașinile virtuale puse în mișcare de către acest sistem.

## Inteligența artificială

Pot fi oare create mașini virtuale tot mai sofisticate, capabile să interacționeze cu utilizatorii cam așa cum ar face-o

---

<sup>84</sup> Promotorul introducerii bazelor teoretice ale programării computerelor la Filosofie a fost Grigore Moisil. Moisil, la cursurile căruia am fost student în 1971-1972, avea încă de atunci viziunea exactă a importanței computerelor, inclusiv a posibilității de a avea acces la un computer personal.

oamenii?<sup>85</sup> Până la un punct, răspunsul este evident afirmativ. Editorul WORD de la catedra noastră răspunde la cereri de ajutor formulate în limbaj natural (în engleză, în cazul nostru). Mașini virtuale, încă destul de scumpe, pot face traduceri mult mai bune decât se crede în mod curent. Campionul mondial la șah a fost învins de un program, iar șahiști de primă mână au mari probleme cu programe care rulează pe computere uzuale, fără viteze sau memorii ieșite din comun.

Exemple ca acelea de mai sus păreau de domeniul basmelor acum un deceniu. Acum ele sunt reale și am putea menționa noi cazuri. Chestiunea pare să fie ceva mai complicată însă. Pentru a diseca resorturile ei interne propun să ne concentrăm doar asupra unora dintre capacitățile minții umane care ar putea fi simulate pe calculator: vederea, memoria, limbajul. Noi oamenii suntem capabili să vedem obiecte și să ne amintim evenimente. Iar ce pare să ne distingă este o rafinată utilizare a limbajului. Pot face asemenea lucruri mașinile virtuale? Și cum ar face-o?

În esență, proiectul inteligenței artificiale constă în simularea unor capacități ale minții umane cu ajutorul unor programe care specifică în mod explicit structura și funcționarea mașinii virtuale. Definiția aceasta este mai îngustă decât cele pe care le veți întâlni în mod uzual și ea urmărește să creeze terenul pentru un contrast între “inteligenta artificială” și “viața artificială”.<sup>86</sup>

Programele la care facem aluzie mai sus sunt adesea scrise în Lisp sau Prolog. Acestea sunt limbaje în sensul în care “limbajul mașină” este un limbaj. Ca și în cazul limbajului

<sup>85</sup> Thomas Dean, James Allen și Yiannis Aloimonos scriu în deschiderea manualului lor de inteligență artificială (*Artificial Intelligence*, Menlo Park, Addison-Wesley, 1995, p.1) că IA (inteligenta artificială) este disciplina care cercetează proiectarea programelor de computer care se comportă în mod inteligent. Aceste programe reacționează în moduri pe care, precizează ei, la animale sau oameni le-am considera inteligente.

<sup>86</sup> Textul lui John A. Eikenberry “Linux & AI-Alife mini-HOWTO”, reprodus în John Purcell (editor), *Doctor LINUX*, ediția a 5-a, Linux Systems Labs, 1997 (vezi și site-ul web [www.lsl.com](http://www.lsl.com)), pp.1577-1591, poate tocmai din pricina caracterului său foarte strict practic m-a convins de utilitatea acestei distincții.



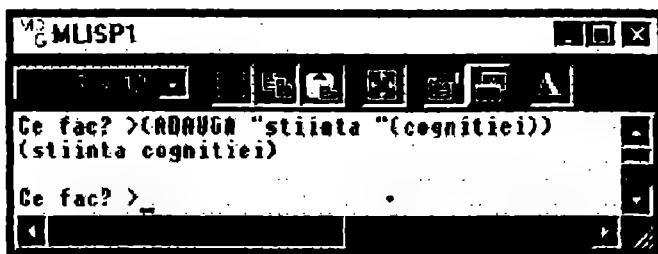
mașină vom folosi o față românească a interpretoarelor scrise de Samuel Kamin pentru Lisp și Prolog, cu scopul de a face viața cititoarei sau cititorului mai ușoară.<sup>87</sup> Mai mult decât atât, unele cuvinte ciudate au primit o versiune inteligibilă. Scopul aici nu este de a învăța pe cineva să programeze în Lisp, ci de a arăta care este stilul de programare în limbajele de tip Lisp.

De fapt, marele pas înainte făcut de Lisp în raport cu limbajul prezentat mai sus constă în faptul că programele Lisp prelucrează liste de simboluri, nu doar numere. Avem oare la nivelul mașinii fizice nevoie de schimbări pentru a prelucra liste de simboluri? Sistemul prezentat până acum nu prelucrează decât numere! Și totuși nu avem nevoie. Amintiți-vă de ideea mașinilor virtuale! Putem să codificăm un simbol precum "a" sub forma unui număr (ca în codul ASCII). Prelucrăm apoi numărul respectiv. Când interacționăm cu mașina virtuală nu este nevoie să știm cum se lucrează cu toate aceste coduri. Programul care face să existe mașina virtuală are grijă de toate aceste lucruri.

Mașina virtuală Lisp extinde efectiv mașina de bază descrisă mai sus. Toate computațiile care puteau fi făcute anterior pot fi făcute în continuare. Pe lângă acestea se poate opera cu liste.

---

<sup>87</sup> Versiunile lui Kamin sunt simplificate în scop didactic. Pentru o inițiere în Lisp puteți folosi cartea Mihaelei Malița, *Antrenamente Lisp*, Editura Universității din București, 1998, 212 p. Vezi bibliografia din carte pentru o îndrumare către alte surse de informare asupra Lisp. Evident, Internetul este, de asemenea, un loc cu multe resurse. Căutați simplu "lisp". Veți găsi, între altele, și cartea celebră despre Lisp a lui G. L. Steele, *CommonLISP. The Language*.



```

MLISP1
Ce fac? >(ADDAUGA "stiinta "(cognitiei))
(stiinta cognitiei)
Ce fac? >

```

În imagine se vede cum numele programului care rulează s-a schimbat (acum este "MLISP1"). Acest program creează mașina Lisp care mă întreabă ce are de făcut. Îi spun că are de adăugat simbolul "stiinta" (atenție la ghilimelele care preced simbolul; sunt importante!) la lista care are un singur element (simbolul "cognitiei"). Ecoul îmi spune care este rezultatul.

Sistemul poate recunoaște dacă un șir de semne este sau nu o listă. Dacă este o listă răspunde cu "A" (de la "Adevăr").

```

Ce fac? >(LISTA? "stiinta)
()

Ce fac? >(SIMBOL? "stiinta)
A

Ce fac? >(NUMAR? 7)
A

Ce fac? >

```

Atenție la semnul de întrebare din "LISTA?". Dacă el lipsește, primim un mesaj de eroare. Dacă o propoziție este falsă, sistemul răspunde cu mesajul "()", care înseamnă în Lisp fie "Fals", fie lista vidă.

Sistemul este capabil să vadă dacă lista este vidă sau nu.

```

Ce fac? >(NUL? "(cognitiei))
()

Ce fac? >(NUL? "() )
A

Ce fac? >

```

O listă are un cap și un corp. Posibilitatea de a le identifica este foarte utilă când se prelucrează listele.

```

Ce fac? >(CAP "(filosofia mintii din
> perspectiva stiintei cognitive))
filosofia

Ce fac? >(COMP "(filosofia mintii din
> perspectiva stiintei cognitive))
(mintii din perspectiva stiintei cognitive)

Ce fac? >

```

Trebuie să fim atenți la scrierea listelor. Simbolurile (elementele listei) sunt separate de blankuri. Dacă trecem pe alt rând, s-ar putea să uităm blankul și să schimbăm conținutul listei.

La ce bun toate acestea? Aceasta este întrebarea pe care o pune conștiința sceptică. Răspunsul este foarte simplu. Chiar "La ce bun toate acestea ?" este o listă. Nu avem decât să concepem dispozitivele virtuale care să o prelucereze în mod adecvat. Tocmai aceasta este provocarea inteligenței artificiale.

Programele care generează dispozitivele virtuale de care vorbeam sunt scrise, cum spuneam, pe o hârtie. Sunt simple texte. De altfel, și mașina virtuală Lisp descrisă mai sus este capabilă să încarce asemenea texte și să ne sugereze ce lucru ciudat este programarea computerelor. Cei care creează dispozitive fizice produc un obiect palpabil. Prin contrast, cei care meșteresc dispozitive virtuale produc ceva impalpabil.

```

Ce fac? >INCARCA
> Textsc
(DEFINITIE SUCCESOR (x) (ADUNA x 1))

SUCCESOR
Ce fac? >(SUCCESOR 9)
10

```

Cuvântul "INCARCA" are un rol special. El face ca sistemul să aștepte numele textului pe care trebuie să-l preia. Textul preluat este afișat și tratat ca un program.

Pentru programe mai sofisticate cu liste trebuie adunată ceva mai multă putere. Deocamdată putem descrie doar în linii mari ce anume face unul dintre programele clasice de inteligență artificială – *Eliza* lui Weizenbaum.

Eliza simulează dialogul cu un psihiatru. Dacă îi spun Elizei “sunt nebun”, ea pune în mișcare un dispozitiv de căutare a unor cuvinte speciale, cum ar fi – de pildă – “nebun”. O dată găsit cuvântul, programul investighează un repertoriu de reacții și produce o replică de genul “de când ești nebun?”.

Eliza nu este un partener de dialog. Efectul programului este acela de a întreține fluxul declarațiilor pacientului și de a provoca o ușurare a tensiunii nervoase prin confesiunile făcute Elizei. Paradoxal, programul cu pricina poate întreține și iluzii cu privire la prelucrarea conștientă a listei de către mașina virtuală (ceea ce nu este adevărat). În orice caz, dacă-l folosiți, atunci simțiți că vă mână o anume curiozitate ciudată. Tot aștepți să vezi ce replică va mai avea Eliza.

## Algoritmi și structuri de date

Tuturor ni se întâmplă probabil să scriem programe, chiar dacă nu ne dăm imediat seama de acest lucru. Motivele care ne împing la așa ceva sunt variate. Iată, de pildă, o situație foarte simplă. La catedra noastră, pe computerul mai vechi era foarte greu să scrii cu litere românești. Ca să nu am surprize de la un program la altul și de acolo la trecerea textelor prin printer, foloseam un sistem simplu de coduri. Pe computerul mai nou din catedră acest sistem nu își are rostul. Editorul Microsoft Word este capabil să lucreze în condiții excelente cu caractere românești. Ce fac însă cu documentele mai vechi? Las acolo codurile vechi? Nu. Scriu un macro cu ajutorul căruia înlocuiesc vechile caractere. Un macro nu este altceva decât un program! Procedez ca în imaginea alăturată.

structuri de date simple

Spre deosebire de schema be

(stimul-react#4ie, mutatis mut

Find and Replace

Find Replace | Go To |

Find what: a#0

Replace with: ă

#0reau pe ecran celule din "1

Indic codurile care urmează să fie înlocuite și caracterul care va lua locul codului vechi. Repet operația până termin toate codurile. Word va scrie automat un program pe care-l pot folosi atunci când vreau să înlocuiesc asemenea coduri. Programul acesta este – ca orice program despre care am vorbit până acum – un text, așa cum se vede și din listingul de mai jos.

Pot să tipăresc acest program. S-ar putea să am dificultăți cu citirea sa, dar aş putea să-i dau o lectură. Pot să schimb textul dacă vreau. De scris nu am scris însă mare lucru din acest text. Am indicat doar *șirul de acțiuni pe care trebuie să le execute programul*, precum și ce anume trebuie să transforme aceste acțiuni.

```

ConvRo Macro
Macro recorded 11/04/99 by Radu M. Solcan

Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormat
With Selection.Find
    .Text = ""
    .Replacement.Text = ""
    .Forward = True
    .Wrap = wdFindContinue
    .Format = False
    .MatchCase = False
    .MatchWholeWord = False
    .MatchWildcards = False
    .MatchSoundsLike = False
    .MatchAllWordForms = False
End With
Selection.Find.ClearFormatting
Selection.Find.Replacement.ClearFormat
With Selection.Find
    .Text = "a#0"
    .Replacement.Text = ChrW(259)

```

Dacă rețin exclusiv aspectul indicării șirului de acțiuni pe care trebuie să le execute programul, atunci putem spune că eu am indicat *algoritmul*. Mai precis am indicat o parte din algoritmul programului.

Dacă ne uităm, pe de altă parte, la ceea ce am scris când am fost întrebat ce anume trebuie înlocuit și cu ce, atunci putem spune că am precizat care sunt *datele* specifice acestui program.

Sunt multe structuri de date și nu ne propunem să le trecem pe toate în revistă aici. Ceea ce vom oferi în continuare sunt doar exemple (utile pentru dezvoltările ulterioare).

Să revenim puțin la adresele din memorie în care operatorul pune ceea ce acum putem denumi date. Dacă datele nu se schimbă de-a lungul întregului proces, spunem că folosim o constantă.

Dacă datele de la o adresă sunt schimbate (sau pot fi, în principiu schimbate) pe parcursul procesului, atunci spunem că folosim o variabilă. Am văzut deja din discuția despre fața umană a limbajului mașină că o variabilă primește un nume cu ajutorul căruia ne referim la ea.

Constantele și variabilele trebuie înțelese din perspectivă procesuală: criteriul luat în considerare este schimbarea sau neschimbarea conținutului unei celule de memorie.

Vom ignora pentru moment constantele și ne vom concentra atenția asupra variabilelor. Variabilele pot fi de diverse tipuri. Stocarea în memorie a datelor conținute de variabile se face mai ușor (și mai eficient) dacă tipul variabilei este definit explicit. De asemenea, prelucrarea datelor respective poate fi orientată, adaptată în funcție de tipul de date prelucrate.

Clasificarea tipurilor de date este destul de complicată și trecere în revistă completă depășește cadrul asumat aici.<sup>88</sup> Ne vom opri, de aceea, pentru început, asupra câtorva tipuri simple de date și vom examina apoi, separat, structuri de date complexe.

Datele zise booleene cuprind exact două elemente, pe care le vom boteza "fals" și, respectiv, "adevărat" (vezi "A" și "( )" în Lisp-ul românizat de mai sus). Acest tip de date este foarte important. Este tipul rezultatului examinării unei declarații, a unei afirmații. Dacă afirm, de pildă, că " $2+2=5$ ", atunci declarația mea este falsă.

Un alt tip de date sunt cele pe care le vom boteza "întreg". Datele de acest tip pot fi adunate, înmulțite ș.a.m.d.

Prin contrast, datele de tip "caracter" pot doar fi ordonate, dar nu adunate. Dacă o variabilă este de tip caracter, atunci dacă ea conține pe "2", acesta nu este un număr pe care să-l putem aduna, ci doar cifra, un caracter, aidoma unei litere pe care o folosim pentru a nota de obicei cuvinte.

---

<sup>88</sup> Pentru o clasificare a tipurilor de date și o prezentare a tipurilor simple de date, vezi, de exemplu, Kathleen Jensen, Niklaus Wirth, *Pascal*, ed. a 3-a, New York, Springer, cap. 2. Noi am folosit traducerea în limba rusă publicată la Moscova de editura Финансы и статистика în 1989.

Structurile de date pot fi tot mai complicate. Un tip de date complex cu care suntem familiarizați și din experiența vieții obișnuite este tabelul. Un tabel are rânduri și coloane.

Când colectăm date empirice ne slujim adesea de tabele: în care coloanele sunt variabilele (în sensul metodologiei de cercetare experimentală și al statisticii), iar pe rânduri sunt consemnate observațiile, înregistrările, constatările noastre.

Mulțimile sunt, din punctul de vedere al științei computerelor, tot un tip de date.

Tipurile de date complexe pot fi privite ca fiind compuse din tipuri de date mai simple. În felul acesta ia naștere o ierarhie de structuri de date<sup>89</sup>. Cu alte cuvinte, pornind de la tipuri de date de bază, se pot construi noi tipuri de date, iar din acestea alte tipuri de date ș.a.m.d.

Structuri de date poate mai ciudate, la prima vedere, sunt cele de genul stivelor. Numele acestei structuri de date, "stivă", ne indică foarte limpede care este și metafora avută în vedere. Datele din care este compusă stiva sunt structurate aidoma unei stive de lemne sau de farfurii, sau de orice alte obiecte. Desigur, aceasta este o metaforă și nu trebuie să ne gândim că așa este realizată această structură de date și în computer.

Important este că o stivă permite două operații: cea botezată "bag" este operația de "stivuire" a datelor unele peste altele; "scot" este operația de extragere. Caracteristic stivei este faptul că "ultimul băgat este primul scos".<sup>90</sup>

O altă structură de date este lista (pe care am întâlnit-o deja în Lisp). Cuvintele scrise pe această pagină sunt, în fond, liste de caractere (de litere). Avantajul definirii unei astfel de structuri ar consta în posibilitatea de a avea la dispoziție conceptele specifice structurii. De exemplu, ideea de "sublistă" este ușor de definit; iar,

---

<sup>89</sup> Cf. Niklaus Wirth, *Algorithms and data structures* [am folosit trad. rusă (Moscova: Mir, 1989)], p. 24.

<sup>90</sup> Este interesant să ilustrăm în acest punct folosirea abstracției ca o unealtă de construit (stive, în acest caz).



concret, articolul hotărât, să zicem, din limba română, ne apare ca sublistă în listele care sunt cuvintele.

De asemenea, structurile de date bine construite ușurează operații cum ar fi sortarea sau căutarea de date. A căuta într-un vraf de date nestructurat ar fi un lucru dificil chiar și pentru computer.

Programele, conform unei caracterizări mai vechi a lui Niklaus Wirth pot fi definite prin următoarea ecuație: structura de date+algoritm = program<sup>91</sup>. Un program spune ce structură au datele care vor fi procesate și care sunt pașii urmați în procesul de prelucrare a lor.

Așa cum am vorbit despre o mașină imaginară, abstractă tot așa putem trata și componentele unui program. *Abstracție*, din perspectiva programelor, înseamnă ascunderea detaliilor legate de mașina concretă pe care vor fi create structuri de date și aplicați pașii unui algoritm.

O stivă, de exemplu, poate fi prezentată ca o structură abstractă de date. Facem abstracție de lucruri cum sunt limbajul în care este scris programul, sistemul de operare sau mașina pe care va rula programul. Iată practic cum ar arăta, în esență, abstracția unei stive după Elliot B. Kaufman.<sup>92</sup>

*Elementele unei stive sunt date de același tip.*

*Structura este determinată de faptul că elementele sunt plasate în ordinea introducerii și doar ultimul element poate fi extras sau copiat și examinat.*

---

<sup>91</sup> Versiunea inițială a cărții lui Wirth despre algoritmi și structuri de date, cea cu exemplele scrise în limbajul Pascal, avea titlul următor: Algorithms + Data Structures = Programs.

<sup>92</sup> E. B. Kaufman, *Pascal - Problem Solving and Program Design*, Reading, Addison-Wesley, 1992, pp.706-707.

*Operații definibile: crearea unei stive vide; adăugarea unui element (succesul sau insuccesul operației este înregistrat de o variabilă de tip boolean - la fel procedându-se și în cazul operațiilor următoare); extragerea (dacă stiva nu este vidă) a ultimei valori adăugate; copierea ultimei valori adăugate stivei.*

*Funcții definibile sunt: o funcție care ia valoarea "adevărat" dacă stiva este vidă; o funcție care ia valoarea "adevărat" dacă stiva este plină (cu alte cuvinte, dacă a fost plasat în stivă numărul maxim de elemente admis).*

O structură de date de tip stivă este foarte potrivită pentru a rezolva, de pildă, problema scrierii literelor dintr-un cuvânt în ordine inversă. Dacă punem cuvântul "stiva" într-o stivă, atunci ultimul element adăugat în stivă este "a" și tot el va fi primul extras ș.a.m.d. Rezultatul este, evident, "avits".

## Funcțiile și recursia

Dacă ați reținut definiția dată anterior conceptului de funcție, atunci ați observat că fiecare dispozitiv, fizic sau virtual, nu face altceva decât să calculeze funcții. Chiar și caracterizarea dată aici dispozitivelor urmează îndeaproape structura modului de prezentare a funcțiilor.

Ce se poate face pentru a calcula o nouă funcție? Desigur, s-ar putea construi un nou dispozitiv fizic, dar lucrul acesta are evidente dezavantaje. Pe măsură ce dispozitivele fizice ar deveni tot mai complexe (și întregul sistem mai complicat!) funcționarea ar fi nesigură. Utilizarea dispozitivelor de dirijare permite însă mutarea confruntării cu complexitatea pe terenul dispozitivelor virtuale.

Să zicem că, așa cum se și întâmplă îndeobște, sumatorul fizic adună doar două numere. Cum vom aduna o listă de numere însă? O posibilitate este să construim programul în felul ilustrat în figură.

```

Ce fac? >INCORCA
> neg
(DEFINITIE NU (x) (DACA x "( ) "A))

NU
Ce fac? >INCORCA
> SumaC
(DEFINITIE SUMAC (lista) (COMPUNE
  (ATRIBUIE rezultat 0)
  (CATTIMP (NU (NIL? lista))
    (COMPUNE (ATRIBUIE rezultat
      (ADUNA rezultat (CAR lista))
      (ATRIBUIE lista (CDR lista))
      (AFISEAZA rezultat))))))
SUMAC
Ce fac? >(SUMAC "(1 2 3))
1
3
6
()

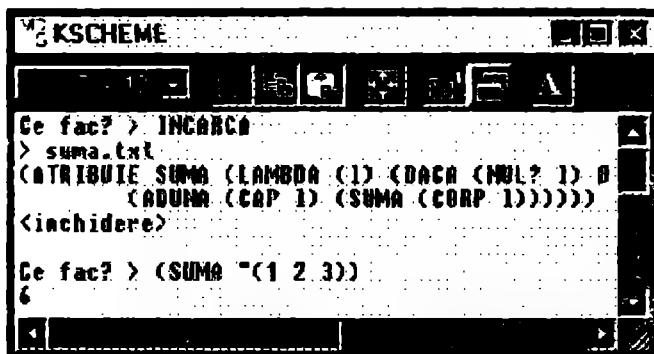
```

Se observă că mai întâi este definită negația. Apoi este definită funcția care calculează suma elementelor unei liste. Algoritmul folosit la calcularea acestei funcții ar trebui să transpară din program. În esență, este vorba despre crearea unei variabile rezultat (în care se pune 0 la început) și de adunarea primului element din listă urmată de tăierea acestuia de pe listă până nu mai rămâne nici un element (observați șirul de rezultate obținute astfel, precum și semnul pentru lista vidă "( )" pus atunci când este falsă condiția ca lista să conțină elemente). Litera "C" din numele funcției SUMAC atrage atenția asupra verificării condiției ca lista să nu fie vidă.

Nu am putea recurge însă la abstracția unei funcții și, implicit, la dispozitivul virtual care o calculează? Am vorbi atunci despre acea funcție care calculează suma elementelor unei liste. Aceasta este ideea care stă la baza sistemului lui Alonzo Church denumit lambda-calcul. Un dialect al Lisp, botezat "Scheme", este limbajul de programare mulat pe ideea lambda-calculului.

Înainte de a vedea un mic program în Scheme, să mai examinăm și structura de date folosită mai sus. Pentru a calcula

funcția SUMAC recurgem chiar la listă: eliminăm treptat capul listei până epuizăm elementele (pe care le-am adunat cu conținutul variabilei unde punem rezultatul). Am putea recurge însă și la o stivă. Ar fi ca și cum un om ar lua lista, ar încerca să vadă dacă poate face suma, dacă ar vedea că nu poate ar pune un element deoparte în stivă și ar continua tot așa până când poate face suma; apoi ar reveni și ar lua pe rând elemente din stivă pe



```

KScheme
Ge fac? > INCARCA
> suma.txt
(ATRIBUIE SUMA (LAMBDA (I) (DACA (NUL? I) 0
  (ADUNA (CAP I) (SUMA (CORP I))))))
<inchidere>
Ge fac? > (SUMA "(1 2 3))
6

```

care le-ar aduna la rezultat.

La prima vedere nu este vorba decât despre inversarea procedurii anterioare. Dacă ne uităm însă la programul Scheme ("K" din "KScheme" vine de la "Kamin", numele celui care a conceput interpretorul de uz didactic căruia noi i-am dat o față românească) din figura de mai jos, vedem ceva care pare foarte ciudat la prima vedere. Cuvântul-cheie "DEFINITIE" a dispărut. Se atribuie doar unei variabile o valoare! Iar valoarea respectivă este o funcție. Este acea funcție (acesta este sensul lui "LAMBDA") care are un argument. Dacă argumentul este o listă vidă, atunci funcția are ca rezultat pe 0; dacă lista nu este vidă se spune ceva și mai ciudat: se va aduna capul listei cu (!!!) ceea ce rezultă în urma aplicării conținutului variabilei în care se află funcția construită cu LAMBDA la corpul listei. Cu alte cuvinte, se adună capul cu suma elementelor care formează corpul.

Cum de nu ne învârtim în cerc? Se procedează, de fapt, cum am sugerat deja: capul listei este pus în stivă și se încearcă

însurarea elementelor corpului. Dacă nu este vid corpul, evident acest lucru nu se poate face cu un dispozitiv ca "ADUNA", care nu face decât suma a două elemente. Se pune iar capul în stivă. Se continuă până când corpul este vid. Atunci se poate face suma elementelor listei! Aceasta este 0. Se pot lua acum elementele puse la păstrare în stivă și problema se rezolvă adunându-le treptat cu sume de elemente. Da, da! Sume de elemente din liste. Ați citit bine. Sumele de elemente care nu puteau fi calculate pot fi acum calculate treptat grație modului în care sunt dispuse elementele în stivă. În exemplul de mai sus, se calculează mai întâi (ADUNA 3 (SUMA (CORP "(3))), deoarece "(SUMA (CORP "(3)))" are valoarea 0; apoi (ADUNA 2 (SUMA (CORP "(2 3))), care are valoarea 5, și, în sfârșit, (ADUNA 1 (SUMA (CORP "(1 2 3))), care conduce la rezultatul final.

Modul acesta de dirijare a computației se numește recursie. Recursia este structura de dirijare preferată în Lisp și dialectele acestuia. Caracteristica ei ușor de reținut este autoreferirea.

Care ar fi avantajul recursiei? La urma urmei, ce putem face cu recursia putem face și cu alte structuri de dirijare. După cum se vede însă din exemplele de mai sus, programele recursive sunt mai compacte. Dacă te obișnuiești să le citești, sunt chiar lizibile. Însă cel mai important pare gradul de abstracție mai mare. Detaliile felului în care este realizată recursia la nivelurile inferioare ale sistemului pot să le ignor. Structurile de nivel înalt pot fi foarte diferite de ceea ce se petrece la nivelurile inferioare.

## Proceduri, obiecte și modularitate

Uneori, în diverse puncte ale programului, se repetă același șir de operații de prelucrare a unor date. Aceste porțiuni ale programelor sunt, la rândul lor, ca niște programe. De altfel, constructorii de programe le și testează separat. Ar fi cât se poate de inutil ca secvențele respective să tot fie rescrise în

program. De aici și ideea de a le stoca în memorie în așa fel încât să poată fi apelate atunci când este nevoie de ele. Secvențele respective se numesc proceduri.

Am putea foarte bine să le spunem procedurilor și subprograme, dar denumirea de “procedură” este mai sugestivă: ea indică faptul că programul precizează felul în care sunt prelucrate date. Este vorba deci despre *modul în care se procedează* pentru a obține un rezultat.

În conformitate cu caracterizarea dată de Wirth programelor, programe scrise în limbaje precum Pascal, creația aceluiași Wirth, cuprind obligatoriu, la început un set de declarații cu privire la structurile de date folosite și un set de proceduri. Programul ca atare este apoi o înstructurare cu ajutorul instrucțiunilor de dirijare a datelor și procedurilor.

Dezavantajul construcției de mai sus este acela că datele și procedurile sunt ținute separate. Tehnica agreată mai nou este aceea de a combina într-un “obiect” datele și procedurile.

Dacă folosiți un program mai nou, dintre acelea care plasează butoane pe ecran, atunci este foarte probabil că butonul respectiv este un “obiect”. Obiectul are proprietăți: formă, culoare, poziție pe ecran etc. Obiectului îi sunt asociate proceduri, care în limbajul programării orientate pe obiecte, se numesc metode. Aceasta înseamnă, de fapt, că butonului îi sunt asociate acțiuni. De pildă, dacă apăsăm pe buton se va efectua un calcul (de genul adunării unei liste de numere, de exemplu). Butonul reacționează la evenimente. “Apăsarea” pe buton (de fapt un semnal trimis cu ajutorul mouse-ului) este un asemenea eveniment.

Cum metodele sunt acțiuni care modifică proprietăți ale altor obiecte, ele funcționează și ca moduri de a trimite “mesaje” de la un obiect la altul. Din această perspectivă, programele sunt lumi, sunt sisteme de obiecte care interacționează. De altfel, obiectele au și fost introduse pentru a rezolva probleme legate de simularea de evenimente și procese din lumea reală.

Obiectele dintr-un program sunt introduse sub forma claselor - abstracții ale obiectelor. Obiectele ca atare sunt instanțieri ale unei clase (exemplare din clasa respectivă). Noi clase pot fi construite din clasele existente prin adăugarea de proprietăți sau metode.

Cel mai important limbaj orientat către obiecte este C++, dar astăzi toate limbajele importante au extensii care fac posibilă programarea orientată pe obiecte.

După trecerea în revistă a nucleului arsenalului de unelte ale programării, la urmă, dar nu pe cel din urmă loc în ordinea importanței trebuie menționată modularitatea. Orice program este compus din module. Fiecare modul are intrările și ieșirile sale și nu se poate sări dintr-un modul într-altul. Motivul este deja amintita problemă a generării unor programe-spaghete, dacă asemenea reguli nu sunt respectate.

Programele modularizate au o serie de avantaje. Mai întâi, ele sunt mai ușor de conceput. În abstracția programului sunt identificate sarcinile, ca să spun așa, pe care le duce la îndeplinire programul. Fiecare asemenea sarcină poate fi desfăcută în sarcini secundare și așa mai departe. Modulele vor fi scrise pentru a rezolva asemenea sarcini.

În al doilea rând, programele modularizate sunt mai ușor de reparat. Trebuie identificat modulul de unde provine eroarea. Cum modulul este o entitate separată, el poate fi extras, corectat și testat separat de alte părți ale programului.

În al treilea rând, programele sunt ușor de dezvoltat dacă sunt modularizate. Pot fi adăugate noi module sau pot fi înlocuite unele module mai vechi cu altele îmbunătățite.

De asemenea, nu este de neglijat nici avantajul practic de a putea elabora programe în echipă. Fiecare membru al echipei scrie module, care apoi sunt asamblate în proiectul final.

### **Cum să facem față complexității?**

Arta programării computerelor nu este altceva decât un răspuns rafinat la provocarea complexității. Computerele și

programele de computer sunt cele mai complicate obiecte produse de oameni. Pentru a le realiza este nevoie de unelte speciale puse sub semnul unor principii folositoare oriunde ne confruntăm cu complexitatea.

- **Simplitatea** [care presupune: simplitatea datelor de intrare și ieșire (biți); dispozitive fizice simple (de prelucrare, dirijare și memorare a datelor)];
- **Ierarhia limbajelor** [limbajul mașină și cele apropiate de acesta includ instrucțiuni de prelucrare, structuri de dirijare și structuri de date mulțate după tiparul dispozitivelor fizice, dar limbajele de la niveluri ulterioare ale ierarhiei utilizează structuri care nu au un corespondent direct la nivelul fizic];
- **Mașini virtuale** [care sunt sisteme de dispozitive virtuale, construite cu ajutorul programelor scrise într-unul dintre limbajele din ierarhia de mai sus];
- **Abstracția** [care presupune realizarea de programe pornind de la structuri abstracte, care ascund detaliile implementării pe un sistem sau altul, dar și folosirea abstracțiilor în limbaje ca atare, ca în cazul claselor din programarea orientată pe obiecte];
- **Modularitatea** [care presupune construirea programelor din module, lucru care permite o proiectare, elaborare, reparare și dezvoltare mai lesnicioasă a programelor].

### Ce recunosc ușor filosofi printre limbajele de programare?

Limbajele descrise mai sus se mai numesc uneori și *limbaje procedurale*. Sub o formă sau alta, programele scrise în aceste limbaje includ proceduri prin care este realizată prelucrarea de date. Nu ar fi însă posibil ca aspectul procedural, în conformitate cu ideea abstracției, să fie cu totul ascuns la niveluri care nu sunt în atenția celei sau celui care programează? Acest lucru este posibil, iar în limbajele (zise *neprocedurale*; mai precis, care ascund aspectul procedural) de acest gen nu există, la nivel superior, decât enunțuri (adevărate sau false) și



relații de consecință între enunțuri plus dispozitivul (zis de obicei *motor*) de dirijare a inferențelor. În continuare, ne vom concentra însă asupra a ceea ce poate fi recunoscut imediat pe baza logicii.

Dacă ținem cont de faptul că sub influența uneia dintre aripile școlii analitice logica a fost tratată ca rezervorul de unelte ce vor permite dezlegarea tuturor problemelor filosofice, atunci putem spune că *uneltele cele vechi ale împăratului* și-au găsit un loc printre cele mai noi.

Limbajul programării logice se numește PROLOG. Vom introduce aici conceptele de bază ale acestui limbaj folosind tot o adaptare a unuia dintre interpretoarele concepute de către Samuel Kamin (de aici denumirea de KPROLOG).

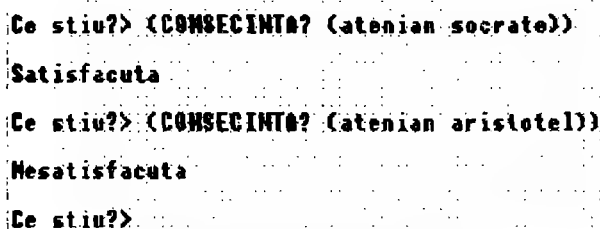
Primul lucru pe care îl poți face cu un sistem PROLOG este să introduci enunțuri despre fapte. Sistemul va trata orice enunț astfel introdus ca pe un enunț adevărat.



Se observă, de la bun început, că sistemul nu mai întreabă ce are de făcut, ci pune o întrebare cu privire la cunoașterea de care urmează să dispună. Pentru început încarcă trei “atomi de cunoaștere”. Cuvântul-cheie “CONSECINTA” introduce un fapt și subliniază o proprietate logică simplă a sistemului: dacă  $p$  aparține sistemului, atunci decurge (este o consecință) a sistemului că  $p$ . Extragerea de cunoaștere din sistem se face cu ajutorul cuvântului-cheie “CONSECINTA?”

(atenție deci, aici “?” nu este un semn separat). Răspunsurile sistemului depind de faptele care-i sunt cunoscute.

În cazul calității de atenian a lui Aristotel sistemul nu poate să tragă nici o concluzie.



```

Ce stiu?> (CONSECINTA? (atenian socrate))
Satisfacuta
Ce stiu?> (CONSECINTA? (atenian aristotel))
Nesatisfacuta
Ce stiu?>

```

Este util pentru obiectivele prezentei introduceri în filosofia minții să distingem cel puțin trei straturi ale cunoașterii, iar sistemul KPROLOG facilitează înțelegerea acestei distincții. Primul strat este cel al *cunoașterii innăscute*. În cazul sistemului KPROLOG acest strat este reprezentat de procedurile ascunse pe care le folosește sistemul pentru a prelucra enunțurile pe care i le furnizăm noi la suprafață. Cum spuneam, acest nivel este ascuns utilizatorului. Avem nevoie însă de acest strat pentru ca sistemul să funcționeze.

Celălalt strat de cunoaștere este alcătuit din *cunoașterea explicită* de care dispune sistemul. În exemplele de mai sus sistemul dispune explicit de fapte cu privire la Socrate și Aristotel.

Ultimul strat de cunoaștere este cel al *cunoașterii implicite*. Sistemul seamănă din acest punct de vedere cu mintea noastră, pentru că cea mai mare parte din timp ne-o petrecem mai degrabă încercând să extragem cunoașterea implicită din cunoașterea explicită. Pentru ca sistemul să dispună de cunoaștere implicită trebuie să-i spunem că anumite fapte decurg din alte fapte (v. exemplul din figură).

```

Ce stiu?> INCARCA
> blatist.txt
(CONSECINTA (blatist X)
  DIN (calator X) (fara_bilet X))
Ce stiu?> (CONSECINTA (calator ion))
Ce stiu?> (CONSECINTA (fara_bilet ion))
Ce stiu?> (CONSECINTA? (blatist ion))

Satisfacuta

```

Prin convenție, literele majuscule sunt folosite în KPROLOG pentru variabile. Perechea de cuvinte-cheie "CONSECINTA ... DIN ..." introduce în sistem cunoașterea necesară pentru a putea extrage ceea ce este implicit în cunoașterea furnizată explicit. În exemplul de mai sus, în faptele "(calator ion)" și "(fara\_bilet ion)" este cuprinsă implicit ideea că Ion merge pe blat, dar pentru a o putea extrage avem nevoie de "(CONSECINTA (blatist X) DIN (calator X) (fara\_bilet X))", care ne spune ceva despre *orice X* care călătorește fără bilet.

Clasicul exemplu cu proprietatea lui Socrate de a fi muritor capătă în KPROLOG forma următoare:

```

Ce stiu?> INCARCA
> muritor.txt
(CONSECINTA (muritor X) DIN (om X))

Ce stiu?> (CONSECINTA? (muritor socrate))

Satisfacuta

```

Aspectul mazochist al sistemului PROLOG ni se dezvăluie însă atunci când vrem să aflăm până și cât fac doi și cu doi. Nu putem face apel la proceduri și nu putem pune deci sistemul să adune numere. Va trebui să punem o întrebare cu privire la ceea ce decurge din ceea ce știe sistemul despre numere.

```

Ce stiu?> (CONSECINTA? (ADUNARE 2 2 4))
Satisfacuta
Ce stiu?> (CONSECINTA? (SCADERE 9 7 2))
Satisfacuta

```

“ADUNARE” și “SCADERE” sunt predicate. Primele lor argumente reprezintă factorii, iar ultimul rezultatul. Acestea sunt singurele predicate primitive aici, așa că restul trebuie construit.

```

Ce stiu?> INCARCA
> mult.txt
(CONSECINTA (INMULTIRE D Y B))
(CONSECINTA (INMULTIRE X Y Z)
  DIN (SCADERE X 1 U)
  (INMULTIRE U Y W) (ADUNARE W Y Z))
Ce stiu?> (CONSECINTA? (INMULTIRE 2 3 6))
Satisfacuta

```

Se observă că recursia este și în PROLOG la loc de cinste.

Sistemul permite, de asemenea, construcții de genul celor din Lisp. De pildă, poți arăta ce înseamnă că un element face parte dintr-o listă.

```

Ce stiu?> INCARCA
> membru.txt
(CONSECINTA (MEMBRU X (ADAUGAT X L)))
(CONSECINTA (MEMBRU X (ADAUGAT Y M))
  DIN (MEMBRU Z M))
Ce stiu?> (CONSECINTA? (MEMBRU 3
> (ADAUGAT 3 4)))

Satisfacuta

Ce stiu?> (CONSECINTA? (MEMBRU 3
> (ADAUGAT 7 (ADAUGAT 3 4))))

Satisfacuta

Ce stiu?> (CONSECINTA? (MEMBRU stiinta
> (ADAUGAT stiinta cognitive)))

Satisfacuta

```

Listele sunt construite prin adăugare de elemente. Calitatea de membru pe o listă este definită recursiv. Calea prelucrării de texte este astfel deschisă și nu este de mirare că PROLOG este cel mai important concurent al Lisp-ului printre limbajele destinate scrierii de programe în inteligența artificială. Avantajul PROLOG este stratul ascuns de proceduri care permit optimizarea căutărilor printre cunoștințele de care dispune sistemul și extragerea de cunoaștere implicită din cunoașterea explicită.

Cu aceasta s-a terminat și incursiunea prin lumea mașinilor virtuale generate cu ajutorul interpretoarelor lui Kamin, așa că putem spune “LaRevedere” sistemului, pentru a-l închide.



Observați pe bara de sus cuvântul “Finished”. Rularea programului a fost încheiată.

## Sistemele expert

Prin anii 1950 pionierii inteligenței artificiale sperau să poată crea sisteme de generale de rezolvare de probleme.<sup>93</sup> Cu timpul, cercetările s-au concentrat asupra *sistemelor expert*. Un sistem expert nu dispune de o cunoaștere generală, ci este construit pentru a rezolva o clasă limitată de probleme.

Sistemele expert au fost primele programe de inteligență artificială produse în scopuri comerciale, nu doar de dragul cercetării.<sup>94</sup>

Un sistem expert combină o *bază de cunoștințe* cu un *motor inferențial* (un mecanism de tras concluzii din baza de cunoștințe). Din punctul acesta de vedere, structura este similară cu aceea văzută în cazul PROLOG. Atâta doar că PROLOG și, cu atât mai mult Lisp, sunt limbaje general utilizabile în inteligența artificială. Există limbaje special dedicate construirii de sisteme expert, cum este, de exemplu, CLIPS.<sup>95</sup>

Pentru a vedea cum funcționează CLIPS să examinăm un program simplu. Programul ca atare este, evident, un text:

(deffacts filosofi

(filosof Platon grec)

(filosof Descartes francez)

(filosof Kant german)

<sup>93</sup> Pentru o introducere accesibilă în ideile clasice ale inteligenței artificiale v. A. M. Andrew, *Artificial Intelligence*, Abacus Press (traducere în rusă la Editura "Mir", 1985), în special capitolul 3.

<sup>94</sup> Vezi exemple istorice în Joseph Giarratano și Gary Riley, *Expert Systems*, Boston, PWS, 1989, pp.19-20. Acolo puteți citi și despre modul în care compania DEC a ajuns să economisească milioane de dolari folosind sistemul XCON.

<sup>95</sup> CLIPS are marele avantaj de a fi liber accesibil. O variantă mai veche se găsește pe discheta atașată cărții lui Giarratano și Riley, citată mai sus. Versiuni mai noi, inclusiv pentru Windows, se găsesc pe Internet. CLIPS a fost creat la NASA. De aceea, NASA este și locul cel mai potrivit pentru a începe căutarea pe Internet a unei noi versiuni a CLIPS. Adresa de la NASA este <http://www.jsc.nasa.gov/~clips/CLIPS.html>

(filosof Russell englez))

(defrule gaseste-filosof-german

(filosof ?nume german)

=>

(printout t ?nume " este un filosof german." crlf))

Prima parte, cea care începe cu "deffacts" reprezintă construcția unei mici baze de date. A doua parte este definiția unei reguli. Sistemele expert bazate pe CLIPS folosesc reguli, în sensul definit anterior în carte. Partea care precede "=>" introduce un șablon. Declanșarea regulii înseamnă, de fapt, căutarea datelor din baza de date care se potrivesc cu șablonul. Partea care urmează după săgeată este acțiunea care este întreprinsă. Variabilele din CLIPS încep cu "?". Programul este încărcat și rulat în felul arătat în prima figură.

```
CLIPS> (load ".\\ExCurs.clp")
$*
CLIPS> (reset)
CLIPS> (run)
Kant este un filosof german.
```

Dacă vrem să vedem felul în care sistemul lucrează dăm comanda "(watch all)".

```
CLIPS> (watch all)
CLIPS> (reset)
==> f-0      (initial-fact)
==> f-1      (filosof Platon grec)
==> f-2      (filosof Descartes francez)
==> f-3      (filosof Kant german)
==> Activation 0      gaseste-filosof-german
==> f-4      (filosof Russell englez)
CLIPS> (run)
FIRE 1 gaseste-filosof-german: f-3
Kant este un filosof german.
1 rules fired
```

Acum se observă lista de fapte accesibile sistemului și indicele fiecărui fapt. Regula este *activată* după f-3 pentru că acesta se potrivește cu șablonul. Se vede apoi că regula este *declanșată* și se execută acțiunea specificată (afișarea unui mesaj).

CLIPS este menit să imite felul în care un neuron declanșează impulsuri electrice. De aici *declanșarea* de reguli. Ca și în creier, un fapt nu poate declanșa o regulă decât o singură dată.<sup>96</sup>

Sistemele expert reale cuprind zeci de mii de reguli. În felul acesta este reprezentată cunoașterea de care dispune un expert. Abia când numărul de reguli este atât de mare, devine evidentă și eficiența sau ineficiența procedurilor de declanșare a regulilor.

Să zicem, de exemplu, că sistemul expert cuprinde reguli de genul “ai cutare simptom => ia cutare medicament” și aceste reguli sunt în număr de 50 000. Se introduce în sistem un fapt legat de starea unui pacient. Care regulă va fi declanșată? Calitatea motorului inferențial va conta atunci enorm. Iar faptul că în creierul uman asemenea căutări sunt făcute relativ repede de expert a dat de gândit specialiștilor în inteligența artificială. De aici și efortul de a simula funcționarea creierului pe care l-am semnalat în cazul CLIPS, dar pe care-l vom vedea și în cazul altor sisteme artificiale.

În rezumat, mintea umană a devenit tot mai mult un model. Iar dacă analiza logică antipsihologistă (“vechile unelte ale împăratului”) clama explicit ideea ignorării gândirii umane și orientarea către limbajele ideale, nevoia de a construi sisteme inteligente eficiente i-a împins pe cercetători înapoi către mintea umană.

### **Cunoaștere generală și cunoaștere încorporată în cazuri**

O puternică tradiție filosofică este legată de cultivarea generalității. Pledoaria clasică pentru căutarea unei idei generale de cunoaștere se găsește la Platon. De exemplu, Socrate îi explică lui Teetet în ce fel trebuie pusă întrebarea despre cunoaștere:

<sup>96</sup> Vezi Giarratano și Riley, *Op.cit.*, p.29 pentru detalii.



Τὸ δὲ γ' ἐρωτηθέν, ὦ Θεαίτητε, οὐ τοῦτο ἦν, τίνων ἢ ἐπιστήμη, οὐδὲ ὁπόσαι τινές· οὐ γὰρ ἀριθμῆσαι αὐτάς βουλόμενοι ἤρουμεθα ἀλλὰ γινῶναι ἐπιστήμην αὐτὸ ὅτι ποτ' ἐστίν.<sup>97</sup>

După Platon întrebarea nu privește diferitele cazuri de cunoaștere sau, cu atât mai puțin, enumerarea acestor cazuri. Întrebarea este ce e cunoașterea ca atare.

Platon recurge la același îndemn și în *Menon* 71d – 77b. în contextul discuției despre virtuți. El accentuează ideea că discuția nu este despre felurile virtuți, ci despre ideea generală de virtute.

Problema setei de generalitate s-a pus, după cum am amintit, în contextul discuției despre sistemele expert, și în inteligența artificială. Am observat cum rezolvarea de probleme a fost restrânsă, de la rezolvarea de probleme în general la domenii limitate.

Adesea însă natura dificultăților legate de cunoaștere nu ține atât de domeniul limitat sau nu, ci de caracterul contextual al cunoașterii. Știu, de pildă, să prepar pizza. Aceasta nu înseamnă că “știu să gătesc”; mai precis, nu are sens să generalizez în acest fel. Aș putea folosi cunoașterea de care dispun deja preparând pizza pentru a coace un fel de pâine. Pot să extind deci cunoașterea la cazuri similare. Dar pe “știu să gătesc” nu am cum să-l definesc.

Există un mod de a dezvolta sistemele expert folosind *raționarea bazată pe cazuri*. Janet Kolodner arată că, din perspectiva programării computerelor, un caz este “un fragment de cunoaștere contextualizată, care reprezintă o experiență ce oferă o lecție fundamentală pentru atingerea scopului urmărit de persoana care raționează”.<sup>98</sup> Un caz are, prin urmare, un conținut (lecția pe care o oferă) și un context. Cea de a doua parte, arată Janet Kolodner, servește la construirea *indicilor* care servesc la

<sup>97</sup> Platon *Theaetetus* 146 e7-10.

<sup>98</sup> Janet Kolodner, *Case-Based Reasoning*, San Mateo, Morgan Kaufmann, 1993, p.13.

clasificarea și regăsirea cazurilor (tot așa cum indicii folosiți în biblioteci servesc la găsirea cărților căutate).

Cunoașterea încorporată în cazuri scoate în evidență rolul fragmentelor de cunoaștere, în contrast cu formulările generale. Dacă este să ne referim la moduri analoage de a trata cunoașterea, în afara științei computerelor, exemplul cel mai faimos este cel al sistemului de drept jurisprudențial, dreptul anglo-saxon bazat pe precedente. În acest sistem, un caz nou este decis pornind de la rezolvarea dată în cazurile similare. Nu este nevoie de o lege cu caracter general pentru a lua hotărâri judecătorești.

Sunt destule situații reale, scrie Janet Kolodner, în care metodele algoritmice nu sunt disponibile și singura cunoaștere este cea încorporată în cazuri.<sup>99</sup>

Reprezentarea cazurilor, în cadrul programelor, este o altă problemă căreia cercetătorii trebuie să-i găsească o soluție. De regulă, reprezentarea unui caz cuprinde trei părți: problema, soluția, rezultatul (aplicării soluției).<sup>100</sup> Abilitatea programatoarei sau programatorului constă, pe de o parte, în construirea unei bune reprezentări a cazurilor, iar pe de alta în proiectarea unor metode eficiente de generare a indicilor și de găsire a cazurilor similare.<sup>101</sup>

## Rețele de crezăminte

Problema “mineritului” în depozite uriase de cunoștințe este una extrem de interesantă. Unul dintre sistemele care o rezolvă cu succes vă salută ori de câte ori deschideți Microsoft

---

<sup>99</sup> Janet Kolodner, *Op.cit.*, p.31.

<sup>100</sup> Janet Kolodner, *Op.cit.*, p.147.

<sup>101</sup> Pentru detalii și prezentarea de aplicații ale sistemelor de cazuri v. David B. Leake (ed.), *Case-Based Reasoning. Experiences, Lessons, and Future Directions*, AAAI Press și MIT Press, 1996, 411 p.

Word și lăsați activ sistemul de ajutor (asistentul care poate să fie o agrafă sau chiar o pisică animată).

Dacă-mi construiesc un tabel și încep să tot pun spații albe pentru a alinia textul pe coloane, după un timp, asistentul începe să se agite și să-mi explice că ar trebui să folosesc tabulatoarele pentru aliniere. Nu i-am spus că vreau să aliniez. Iar sistemul nu are dinainte puse explicit în el tot felul de reacții la comportamentele utilizatorilor. Asistentul urmărește, de fapt, acțiunile mele și consecințele lor. Își făurește apoi *crezăminte* cu privire la ceea ce vreau să fac. Compară modul meu de acțiune cu ceea ce știe deja despre modul optim de lucru în Word și-mi atrage atenția dacă felul în care lucrez eu este impropriu.

Pentru a-și forma *crezăminte* asistentul trebuie să “sape” în noianul de date cu privire la acțiunile utilizatorului și consecințele lor. Firma Microsoft, prin departamentul ei de cercetare, pune la dispoziție (accesul este liber la acest program!) sistemul cu ajutorul căruia se pot forma *rețele de crezăminte* (belief networks).<sup>102</sup>

O firmă specializată în rețele de crezăminte este Norsys. O versiune limitată, utilizabilă pentru a învăța lucrul cu rețelele de crezăminte sau pentru a încerca programul, poate fi obținută pe Internet.<sup>103</sup>

O rețea de crezăminte extrage tipare ale legăturilor cauzale între informațiile stocate într-o bază de date. Acest lucru este făcut cu ajutorul unor proceduri bazate pe calculul probabilităților (de aici și denumirea de rețele Bayesiene sau de rețele probabiliste). Evident, aceste proceduri își au sălașul în stratul de “cunoaștere înăscută” de care dispune sistemul și utilizatorul nu este direct în contact cu ele.

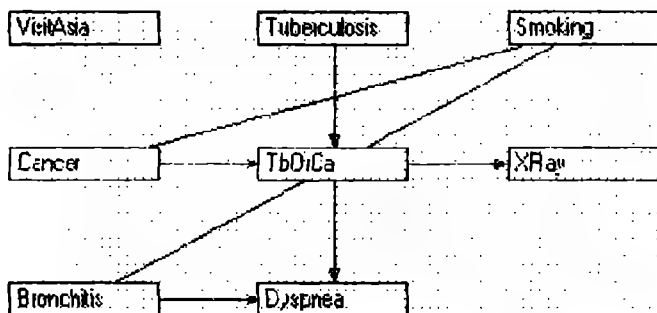
Exemplele clasice de asemenea rețele sunt din domeniul medical. Un medic oferă datele inițiale care sunt introduse într-un tabel. Pe baza acestuia se construiește rețeaua.

---

<sup>102</sup> Programul se numește *Microsoft Belief Networks* și este disponibil pe Internet. Adresa Microsoft (URL-ul, cum se spune tehnic) este <http://www.microsoft.com>.

<sup>103</sup> Adresa pe Internet este <http://www.norsys.com>.

Pentru a construi rețele puteți folosi, cu titlu de evaluare și încercare, programul Belief Network PowerConstructor realizat de Jie Cheng.<sup>104</sup> Exemplul clasic în literatura de specialitate<sup>105</sup> al cauzelor unor boli poate fi lucrat cu setul de date inclus în program și conduce, într-o primă aproximație, la un rezultat de genul celui din figura alăturată.



Se observă că nu există nici o legătură între vizita în Asia a bolnavului și bronșită, de pildă. Între fumat și cancer însă poate exista o legătură.

Ceea ce este fascinant în rețelele de crezăminte este că se poate lucra și cu date calitative, de genul celor pe care le-am pune într-un tabel dintr-o bază de date în care am consemna, să zicem, politicieni, apartenența lor la un partid sau altul, locul de naștere, alte date, precum și acceptarea sau nu a sondajelor de opinie, folosirea pe care le-o dau și așa mai departe. Sistemul ar construi o rețea de crezăminte cu privire la cauzele unui comportament sau a altuia.

Aplicații mai pedestre includ diagnosticarea defectelor printerului atașat calculatorului sau a altor probleme survenite în funcționarea programului.

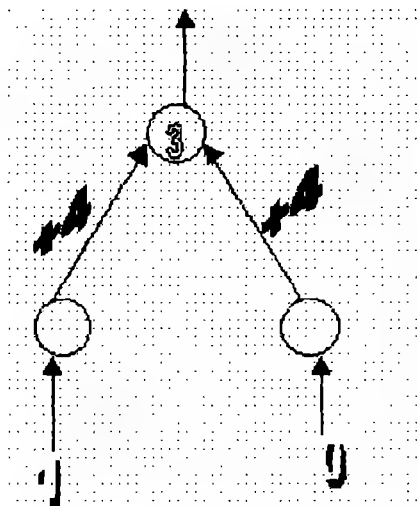
<sup>104</sup> Programul este disponibil apelând pagina de pe Internet a autorului <http://www.cs.ualberta.ca/~jcheng/bnpc.htm>

<sup>105</sup> Steffen Lauritzen și David J. Spiegelhalter, "Local computations with probabilities on graphical structures and their application to expert systems", în *Journal of Royal Statistics Society B*, 50, 2/1988, pp.157-194.

## Viața artificială

Din punctul de vedere care ne interesează aici, cel al studierii minții umane, dintre sistemele artificiale inspirate de natură<sup>106</sup> cele care merită cea mai mare atenție sunt sistemele care imită creierul uman. Imitația aceasta poate fi la nivelul fizic sau la acela al unei mașini virtuale realizate printr-un program adecvat de calculator.

Un prim pas în direcția modului de a funcționa al creierului l-am văzut deja când am prezentat sistemul CLIPS. Modul de a stoca însă informația în CLIPS, dar mai ales felul în care sunt proiectate programele nu este însă o încercare de a imita creierul. Este o încercare de a ajunge la rezultate similare, dar nu neapărat pe aceleași căi – lucru pe care-l putem spune în genere despre inteligența artificială.



În creierul uman sunt circa o sută de miliarde de neuroni. Se pare, de altfel, că și complexitatea neuronilor umani este foarte mare. Neuroni artificiali sunt concepuți, prin contrast, în conformitate cu principiul simplității.

O rețea de neuroni artificiali foarte simplă este prezentată în desenul din prima figură alăturată.<sup>107</sup>

<sup>106</sup> Pentru resursele disponibile pe Internet v. John A. Eikenberry, *Op.cit.*, în special părțile 6 și 7. Un farmec aparte îl prezintă celebrul sistem *Life*, conceput de John Horton Conway, pe care mulți l-au văzut, probabil, sub o formă sau alta.

<sup>107</sup> Figurile sunt adaptate după cele din cartea lui Philip Johnson-Laird, *The Computer and the Mind. An Introduction to Cognitive Science*, Londra, Fontana, 1988, pp.183, 185. O analiză mai amănunțită a funcționării și istoriei lor găsiți în "Biblia de la San Diego" (cartea care a consacrat

Cercurile reprezintă *neuroni*. Săgețile reprezintă *conexiuni* între neuroni. Cifrele de jos (cele cu umbră) reprezintă *intrări*. Cifrele scrise gros și oblic reprezintă *pondera* conexiunii. Ponderile pot fi pozitive (în care caz, neuronul către care se trimite un semnal prin conexiunea respectivă se excită) sau negative (iar, în acest caz, neuronul se inhibă). Cifra scrisă în cercul care reprezintă neuroni este *pragul* de excitare care trebuie depășit pentru ca neuronul să transmită un semnal.

Neuronii funcționează după principiul “totul sau nimic”: dacă pragul este depășit se transmite un semnal (și putem marca acest lucru, în mod convențional, printr-un “1”); dacă pragul nu este depășit, neuronul nu transmite un semnal (sau semnalul este “0”). Excitarea neuronilor se calculează însumând produsele semnalelor de intrare cu ponderile conexiunilor care vin către neuronul respectiv.

O rețea ca aceea de mai sus se numește *rețea de neuroni artificiali* (pe scurt, *rețea neuronală*). Realizarea de dispozitive computaționale cu ajutorul unor rețele neuronale se numește *conexionism*.

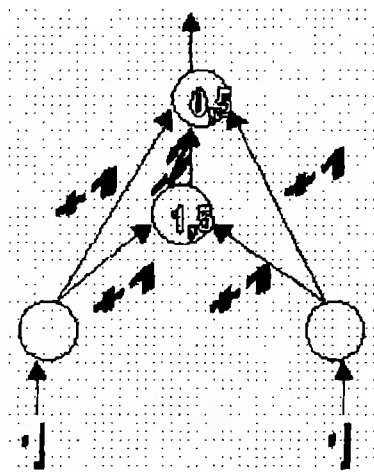
Prin modificarea ponderilor și a pragurilor, rețelele pot fi transformate în așa fel încât să realizeze diferite dispozitive. Acesta este modul de a programa o rețea și este evident că el reprezintă o alternativă la ceea ce se numește *computaționalism*: programarea cu ajutorul codurilor pentru dispozitive și date, așa cum a fost ea descrisă anterior.

Dispozitivele conexioniste realizează și ele computații și de aceea ni se pare mai comod să vorbim despre o *abordare computațională* atât în cazul computaționalismului, cât și al conexionismului.

Rețeaua de mai sus este un dispozitiv care realizează funcția logică “sau”. Există, după cum se vede, două valori la intrare. Combinația folosită cu titlu de exemplu este “1 0”. Alte combinații posibile sunt: “0 1”, “1 1”, “0 0”. Rezultatul este “1” sau “0” în funcție de declanșarea sau nu a neuronului de ieșire din rețea (cel de deasupra). Să verificăm funcționarea pentru valorile de intrare din figură. Neuronul de intrare din dreapta nu transmite nici un semnal și nu contribuie la excitarea neuronului de ieșire din rețea. Cel din

stânga, în schimb, va excita neuronul de ieșire pe o conexiune cu ponderea 4. Pragul de excitare este depășit și neuronul de ieșire emite un semnal (ieșirea este "1"). Se poate verifica ușor că pentru intrările "0 1" și "1 1" ieșirea este "1", iar pentru "0 0" ieșirea este "0".

Am văzut că sistemele computaționale nu pot funcționa fără dispozitive de dirijare condiționată. Or, pentru a realiza asemenea dispozitive trebuie comparate valori. Cu alte cuvinte, cel puțin o funcție XOR (un "sau exclusiv") este necesară. Rețeaua din figura următoare realizează o funcție logică XOR.



Rețeaua aceasta ilustrează rolul conexiunilor negative. În pofida faptului că ambele intrări au valoarea "1", neuronul de ieșire nu se declanșează pentru că valorile de excitare pozitive transmise de la intrări sunt anulate de impulsul primit de la neuronul intermediar. Acest neuron din stratul intermediar, cel cu pragul de excitare de 1,5, în schimb, se declanșează, deoarece suma excitărilor primite de la intrări depășește pragul. Dar pragul nu este depășit atunci când doar una dintre intrări are valoarea "1".

Se observă cum, dacă am modifica ponderea conexiunilor, dispozitivul s-ar schimba. Ieșirile ar putea fi altele. Programarea ieșirilor ar putea fi realizată prin introducerea ponderilor adecvate. Aici intervine o mare deosebire între computaționalism și conexiionism. În cazul computaționalismului programul care conduce la ieșirile dorite este introdus în memoria calculatorului. În cazul rețelei neuronale nu se folosesc programe care să specifice în mod explicit cum se obțin valorile de ieșire adecvate.

Cu alte cuvinte, rețeaua neuronală depășește obiecția curentă la adresa computerelor, conform căreia acestea nu fac decât ceea ce le spun programatorii sau programatoarele să facă.

Rețeaua se adaptează la mediul înconjurător asemenea unei ființe vii.

Majoritatea rețelelor neuronale sunt construite pentru ca să învețe ceva. Unele învață, de pildă, să distingă o mină marină de o stâncă al cărei colț iese din apă. Altele recunosc literele de pe o pagină tipărită. Altele reușesc să prevadă tendința ratelor dobânzilor. Cum ajung la acest rezultat?

O rețea neuronală este tot un computer, făcut din dispozitive de genul celor descrise mai sus. Arhitectura rețelei (genul de dispozitive, structura lor, conexiunile) și ponderile conexiunilor inițiale sunt fie create fizic, dacă rețeaua este un sistem, fie sunt create printr-un program (dacă rețeaua este o mașină virtuală). Toată arta constă apoi în scrierea de algoritmi care permit antrenarea rețelei. Rețeaua primește intrări și produce ieșiri care sunt comparate cu scopul urmărit. Rolul algoritmilor folosiți la antrenarea rețelei este să genereze ajustările necesare ale ponderilor conexiunilor.

Ponderile conexiunilor sunt deosebit de importante. Dacă ne gândim la rețea ca la o mașină virtuală, atunci strategia antrenării este următoarea: se rulează programul cu un set de ponderi ale conexiunilor; apoi conexiunile sunt modificate și programul este rulat din nou; procesul acesta este repetat până este atins scopul urmărit prin scrierea programului.<sup>108</sup>

Pentru ca procesul de antrenare să poată decurge efectiv algoritmi de ajustare a ponderilor trebuie să preia ieșirile, să producă un rezultat în urma calculării unei funcții și să folosească acest rezultat ca pe un *feedback*, care este introdus într-unul dintre straturile de neuroni (fie cei de intrare, fie neuronii ziși ascunși – de genul neuronului cu prag 1,5 în desenul de mai sus).

---

<sup>108</sup> Cf., pentru procesul de antrenare a rețelei, Valluru B. Rao și Hayagriva V. Rao, C++ *Neural Networks and Fuzzy Logic*, New York, MIS:Press, 1993, pp.5-6. Așa cum reiese și din titlul acestei cărți, limbajul C++ este utilizat pentru scrierea programelor care realizează rețele neuronale. Puterea legendară a limbajului C++ este, după cum se vede, necesară pentru acest gen de programe.



O rețea poate să învețe supravegheat sau nesupravegheat<sup>109</sup>. Când procesul de ajustare a ponderilor conexiunilor decurge fără criterii externe care să constrângă ieșirea din rețea (învățare nesupravegheată), se mai spune că rețeaua se *autoorganizează*.

Ideea de a imita funcționarea neuronilor există de relativ mult timp.<sup>110</sup> Handicapul l-a reprezentat multă vreme incapacitatea de a crea dispozitive gen XOR. O dată cu completarea corespunzătoare a arsenalului de dispozitive, rețelele neuronale au ajuns în mare vogă.<sup>111</sup> O parte a succesului de care se bucură în momentul de față rezidă în capacitatea lor de a furniza replici ale ideilor din inteligența artificială.<sup>112</sup> Sistemele cu reguli, de genul celor amintite în cazul sistemelor expert, pot fi realizate și cu rețele neuronale. În plus însă, rețelele sunt mai comod de construit decât un program care ar trebui să producă "manual" rezultate pe care o rețea le obține prin antrenare într-o noapte.

Prezentarea ideii de viață artificială, fie și numai din punctul de vedere al minții, nu ar fi totuși completă dacă nu am aminti de algoritmi genetici. Un algoritm genetic, de pildă, poate ilumina ideea de agent rațional, atât de importantă în științele sociale.

---

<sup>109</sup> *Ibidem*.

<sup>110</sup> Ca și inteligența artificială, modelele neurale există încă din anii 1950. Ideea este anticipată chiar și în anii 1940, în studiul clasic al lui McCulloch și Pitts despre neuroni și calcul logic (1943). Cibernetica este, de asemenea, ca să zicem așa, versiunea mai veche a conexiunismului.

<sup>111</sup> *Parallel Distributed Processing: Explorations in the Microstructures of Cognition* (1986) este cartea care a marcat ascensiunea formidabilă a ideii de rețea neuronală. Ea mai este numită popular și "Biblia de la San Diego", pentru că a fost editată de un grup de cercetători de la Universitatea din San Diego (SUA), condus de David Rumelhart și Jay McClelland. Miezul cărții îl constituie descoperirea unui algoritm eficient pentru antrenarea rețelelor prin retropropagarea erorilor din strat în strat.

<sup>112</sup> Vezi, de exemplu, cercetările lui Saito și Nakano consacrate sistemelor expert conexiuniste (cf. Matthew Zeidenberg, *Neural Network Models in Artificial Intelligence*, New York, Ellis Horwood, 1990, pp. 131-134).

Toți cei care s-au familiarizat cu filosofia economiei, cu filosofia politică sau teoria instituțiilor sociale au citit, desigur, despre "dilema prizonierului". Foarte pe scurt ea poate fi rezumată printr-o poveste de acest gen: doi hoți comit împreună o spargere importantă; poliția îi arestează, dar nu are dovezi decât în sprijinul unor acuzații minore; poliția îi separă pe cei doi suspecti, pentru a-i împiedica să se pună de acord; fiecare dintre suspecti are de ales între a colabora cu poliția (și a scăpa astfel cu o pedeapsă foarte ușoară) și a refuza să furnizeze poliției dovezi despre spargerea cea mare. Dacă nu colaborează cu poliția și celălalt nu colaborează nici el, atunci suspectul scapă cu o pedeapsă mică. Problema pentru fiecare dintre suspecti este însă că nu știe ce va face celălalt. Va fi loial? Atunci ambii suspecti scapă ieftin. Dacă nu este loial, atunci primește o pedeapsă grea. Ce este rațional să faci? Dacă judeci rațional, atunci mărturisești. Te ferești de temnița grea pe care ți-ar aduce-o mărturia celui alt cuplată cu tăcerea ta. Rezultatul este însă prost: o acțiune coordonată ar fi făcut ca ambii suspecti să scape cu pedepse mai mici.

Dacă problema de mai sus este pusă în mod repetat, atunci toată arta este să descoperi un partener loial. Este, evident, nu doar o problemă a hoților, ci și a noastră, a celor cu mai mare teamă de lege. Fiecare, în familie, în afaceri, în parlament avem nevoie de parteneri loiali. Robert Axelrod a cercetat asemenea situații în care dilema prizonierului se pune în mod repetat. El a investigat felul în care indivizii descoperă o regulă de alegere a unui partener loial. Regula folosită trebuie să evolueze în așa fel încât să permită atingerea scopului vizat. În esență, cea mai bună regulă este să fii loial în primul tur, iar apoi să faci partenerului ce ți-a făcut el sau ea în turul anterior.<sup>113</sup>

Rezultate de acest gen au fost dezvoltate sub forma unor "rețele genetice", în care nu ajustarea ponderilor conexiunilor este mijlocul prin care rețeaua se apropie de atingerea scopului, ci ajustarea regulilor.

<sup>113</sup> Pentru mai multe detalii și legătura cu rețelele neurale, semnalată în continuare, vezi Matthew Zeidenberg, *Neural Network Models in Artificial Intelligence*, New York, Ellis Horwood, 1990, pp.92-93.

Schema generală a unui algoritm evoluționist, după D. Fogel,<sup>114</sup> este următoarea: inițializarea procesului prin introducerea unei serii de soluții (care poate fi făcută prin eșantionare aleatoare); calcularea unui indice de performanță al soluțiilor; varierea și selecția soluțiilor. Procesul acesta de selecție a soluțiilor este aplicat în orice situație în care este vorba despre alegerea unui mod optim de a proceda. Iar, după Fogel, programarea evoluționistă oferă și cele mai bune rezolvări pentru problemele din lumea reală.

Cel mai mare avantaj al programării evoluționiste, susține Fogel, este oferirea de soluții acolo unde nu se cunosc (de către oameni) rezolvări ale problemei date. Este o metodă de a găsi metode de rezolvare a problemelor (o veche ambiție a filosofiei științei).

Calculațiile evoluționiste au fost aplicate, de asemenea, la antrenarea de rețele neuronale.<sup>115</sup> Se combină astfel, susține P. Angeline, flexibilitatea programării evoluționiste cu organizarea prezentă în cadrul unei rețele neuronale.

Programarea evoluționistă are însă o presuposiție care o face să treacă dincolo de orizontul rețelelor neuronale. După cum explică David B. Fogel și L. J. Fogel, inteligența artificială (în care ei includ atât computaționalismul, cât și conexiunismul) s-a centrat pe imitarea comportamentului uman și a

---

<sup>114</sup> D. B. Fogel, "The Advantages of Evolutionary Computation," *Bio-Computing and Emergent Computation 1997*, D. Lundh, B. Olsson, and A. Narayanan (eds.), Sköve, Sweden, Singapore, World Scientific Press, 1997 pp. 1-11. Disponibil pe Internet la <http://www.natural-selection.com/people/dbf/docs/bcec97.zip>

<sup>115</sup> Cf. Angeline PJ, "Multiple Interacting Programs: A Representation for Evolving Complex Behaviors," *Cybernetics and Systems*, 29 (8), 1998, pp.779-806. Disponibil pe Internet la <http://www.natural-selection.com/people/pja/docs/mips3.zip>

neurofiziologiei creierului uman.<sup>116</sup> Presupoziția programării evoluționiste este că, pentru a modela efectiv inteligența, demersul trebuie centrat pe emularea proceselor evolutive. În rezumat, programarea evoluționistă este poate într-un sens mai deplin al cuvântului viață artificială.<sup>117</sup>

## A fi în lume

După cum se vede deja din prezentarea ideii de viață artificială, programarea sistemelor artificiale are în structura ei fibre filosofice profunde. Urzeala întregului sistem este, până la urmă, filosofică. Ce înseamnă cunoaștere? Putem avea o cunoaștere generală? Ce înseamnă o reprezentare a realității? Ce sunt metodele de rezolvare a problemelor? La aceste întrebări și la altele, în bună tradiție filosofică, se poate face tentativa de a pune în discuție presuposițiile întrebărilor respective.

Sistemele conexioniste, de pildă, deși ocolesc ceea ce este poate capcana furnizării explicite a întregului program, presupun totuși existența unor *reprezentări* și a unor *comutații*. Sistemul care deosebește, de exemplu, o stâncă de o mină marină își formează, în urma antrenării, o reprezentare a minei. Iar rețeaua XOR realizează o comutație. Nu există o procesare centralizată a datelor, dar de la un subsistem la altul al unei rețele complexe se trece după ce s-a realizat o comutație.

Filosofic este foarte interesant să amintim că existența unei reprezentări poate fi interpretată ca presupunând o punere

---

<sup>116</sup> D. B. Fogel și L. J. Fogel, "Evolution and Computational Intelligence," *Proceedings of the 1995 IEEE International Conference on Neural Networks*, Perth, Australia, Special Session on Computational Intelligence, IEEE Press, 1995.

Disponibil pe Internet la

<http://www.natural-selection.com/people/dbf/docs/perth2.zip>

<sup>117</sup> Pentru o sinteză istorică v. P. J. Angeline PJ (1998) "A Historical Perspective on the Evolution of Executable Structures," *Fundamenta Informaticae*, 36, 1-4 (august) 1998, pp. 179-195. Disponibil pe Internet la <http://www.natural-selection.com/people/pja/docs/gphist.zip>

Articolul conține și o bibliografie substanțială.

în afara lumii, cam așa cum o fotografie a unei pietre are de a face cu piatra doar în sensul că este o reprezentare a ei, dar nu în sensul în care un om ar avea de a face cu pietrele. Asemenea presupoziii sunt ținta criticilor filosofice. Martin Heidegger a introdus în filosofie ideea de "a-fi-în-lume". În stilul său caracteristic, el scrie:

*In-Sein (...) meint eine Seinsverfassung des Daseins und ist ein E x i s t e n z i a l. Dann kann damit aber nicht gedacht werden an das Vorhandensein eines Körperdinges (Menschenleib) ...*

*... In-Sein ist demnach der formale existenziale Ausdruck des Seins des Daseins, das die wesenhafte Verfassung des In-der-Welt-seins hat.*<sup>118</sup>

În comentariul său la *Sein und Zeit*, Hubert L. Dreyfus consideră că:

*... Heidegger pare a sugera că posedarea unui corp nu ține de structura esențială a Dasein-ului (...) Din generalitatea modului de a fi o activitate prin excelență de autointerpretare al Dasein-ului decurge, fără îndoială, că Dasein-ului nu este în mod necesar întrupat. (...)*

*Heidegger ne oferă o descriere lămuritoare a modurilor diferite în care obiectele și oamenii sunt în lume. (...) Deobicei nu luăm în seamă rosturile diferitelor sensuri ale multora dintre prepozițiile și expresiile noastre tocmai pentru că le folosim într-un mod atât de transparent. Dacă ne detașăm însă de acest uzaj și reflectăm la înțelesul unei prepoziții cum ar fi "în", atunci primul sens care ne vine în minte este cel categorial, cel de incluziune fizică.*

<sup>118</sup> Martin Heidegger, *Sein und Zeit*, ediția a 4-a, Halle, Max Niemeyer. 1935, p.54.

*Când cineva ne atrage atenția că "în" are și un sens existențial care exprimă implicarea, precum atunci când spunem că suntem în-drăgostiți, că suntem în problemă sau că suntem în lumea teatrului, tindem să tratăm toate aceste utilizări ca pe o extindere metaforică a sensului de incluziune fizică. Este exact lucrul la care ne-am aștepta dacă Heidegger are dreptate atunci când spune că Dasein-ul se (rău)interpretează pe sine în termenii obiectelor cu care are de a face.*<sup>119</sup>

Dreyfus, avansând în reconstrucția sa a gândului lui Heidegger, pune în evidență și viziunea diferită despre cunoaștere care este asociată cu ființa-în-lume; cunoașterea teoretică pierde prioritatea ei în favoarea unui "know-how" marcat de implicarea în lume. Dreyfus îl aseamănă pe Heidegger cu autori precum Michael Polanyi și Thomas Kuhn.<sup>120</sup>

Asemenea gen de idei, chiar dacă par pur speculative, se regăsesc în presupuzițiile unora dintre modurile de a construi sisteme artificiale inteligente. Un exemplu pare a fi în programare orientarea centrată pe sistemele dinamice. Promotorii sistemelor dinamice critică conexiionismul pentru ceea ce ei socotesc a fi menținerea ideii de reprezentare și a computațiilor, ca în computaționalism.

"Sistem dinamic" este poate pentru mulți o sintagmă opacă. Tot ce pot să fac este să explic cum aș încerca eu să risipesc ceața. Mai întâi trebuie să fim atenți că un "sistem" presupune cel puțin *stări* și *secvențe* de stări în care se află sistemul. Starea sistemului nu este altceva decât mulțimea valorilor variabilelor care caracterizează sistemul (valori consemnate la un moment dat). Ideea de *sistem dinamic* nu înseamnă altceva decât că stările sistemului se schimbă (după o

<sup>119</sup> Hubert L. Dreyfus, *Being-in-the-World*. A Commentary on Heidegger's *Being and Time*, Division I, Cambridge, Mass., The MIT Press, 1991, p. 41.

<sup>120</sup> *Ibidem*, p. 46.

anumită regulă). Altfel spus, stările ulterioare depind de stările anterioare.<sup>121</sup>

Un sistem dinamic poate fi descris în limbaj matematic. Iar aici ecuațiile joacă un rol-cheie. Nu trebuie să uităm că o ecuație este o propoziție (adevărată sau falsă), astfel încât un sistem de ecuații este ca un text care descrie sistemul. Pentru a sesiza dinamica sistemului este nevoie să fie surprinsă *rata schimbării unei variabile la un moment dat* (în funcție de valorile celorlalte variabile). Acest lucru se face cu ajutorul *ecuațiilor diferențiale*.<sup>122</sup>

Abordarea centrată pe sistemele dinamice a inteligenței subliniază faptul că organismul uman este un sistem dinamic integrat în lume, pentru care distincția obiect-subiect nu are sens.<sup>123</sup> Inteligența, din această perspectivă, este o funcție a sistemului constituit din corp și mediu.

Recursul la reprezentări este respins de Rodney Brooks, profesor și constructor de roboți de la MIT. După Brooks, sistemele inteligente nu trebuie (des)compuse într-un set de dispozitive care procesează în mod independent (reprezentări) și își trimit de la unul la altul reprezentări. Sistemele inteligente se compun mai degrabă din dispozitive care interacționează direct cu lumea (printr-un ciclu percepție-acțiune) și funcționează în

<sup>121</sup> Am găsit foarte util pentru a scrie aceste rânduri "A Brief Dynamical-Systems Glossary" anexat de Edward N. Lorenz la cartea sa *The Essence of Chaos*, University of Washington Press, 1993, pp.205-213.

<sup>122</sup> Pentru cei care s-ar putea gândi că aceste concepte ale analizei matematice sunt închise doar în cărți destinate celor care ocupă cu matematica sau științele naturii, ar trebui precizat că există excelente manuale de analiză pentru umaniști (v., de pildă, James Callahan et al., *Calculus in Context*, New York, Freeman and Company, 1993). Lumea ratei schimbării variabilelor poate fi cât se poate de fascinantă pentru oricine.

<sup>123</sup> Vezi despre sistemele dinamice articolul doamnei Margaret A. Boden. "Artificial Intelligence", în *Routledge Encyclopedia of Philosophy*, ediția pe CD, versiunea 1.0. De asemenea, vezi T. Van Gelder, "What is Cognition, if not Computation?", în *Journal of Philosophy*, 91/1995 (apud Boden, *Op.cit.*).

paralel. Noțiunile de sistem central și sistem periferic se evaporă în acest fel.<sup>124</sup>

Robotul Cog, construit de Brooks la MIT, denotă o preocupare a celor care l-au proiectat în primul rând pentru percepție și acțiune.<sup>125</sup> Lucruri simple, precum joaca sau deplasarea de obiecte, sunt minuțios meșterite de Rodney Brooks și echipa sa. Impresia de ființă vie este într-adevăr mai puternică la acest robot al cărui comportament amintește de acțiunile unui animal de casă. Accentul nu mai este pus în construcția robotului pe sisteme centrale de procesare menite să imite rațiunea umană, ci pe "a-fi-în-lume".

Brooks menționează posibila analogie cu Heidegger, dar refuză să dea curs oricărui *parti pris* filosofic. Fiecare este liber, după Brooks, să folosească aceste sisteme artificiale drept sprijin pentru o poziție filosofică sau alta.

### Sistemele artificiale: de la confruntări la rezolvări practice

Înainte de a rezuma pașii făcuți în examinarea sistemelor artificiale, ar trebui precizat că – în ciuda chiar a confruntărilor dintre școli și a tonului polemic adesea aspru – sistemele artificiale prezentate pot fi hibridate.

Larry R. Medsker menționează o serie de sisteme inteligente hibride.<sup>126</sup> Un astfel de sistem poate să combine, de pildă, o rețea neuronală cu un sistem expert. O combinație adesea prezentă este cea dintre sisteme fuzzy (care utilizează logici neclasice cu valori nuanțate, nu doar tranșantele "adevărat", "fals") și rețele neuronale. Sistemele expert pot fi și ele combinate cu o logică vagă.<sup>127</sup> Programele evoluționiste

<sup>124</sup> Rodney Brooks, "Intelligence without representation", *Artificial Intelligence*, 47/1991, pp.139-159.

<sup>125</sup> Robotul Cog poate fi văzut pe CD-urile editate de B. Levy și E. Servan-Schreiber, *Les Secrets de l'intelligence*, Paris, Hypermedia.

<sup>126</sup> Cf. Larry R. Medsker, *Hybrid Intelligent Systems*. Boston, Kluwer, 1995, 300 p.

<sup>127</sup> Există, de exemplu, FuzzyCLIPS, o extensie a sistemului CLIPS, o unealtă dedicată special construirii de astfel de sisteme.



sunt adesea combinate cu rețele neuronale și sisteme fuzzy. O posibilitate mai nouă o reprezintă combinarea de sisteme bazate pe cazuri cu alte sisteme. Ideea aceasta apare în mod natural, deoarece aceste sisteme sunt în aceeași categorie cu sistemele expert.

Contrastul pe care l-aș creiona ar fi de aceea unul între două tipuri de interese cognitive și două tipuri de metodologii.

- **Sistemele artificiale ca sisteme logice [în acest caz interesul cade pe obținerea de rezultate cu valoare universală; problema este, pe de o parte, că aceste rezultate ar putea fi obținute însă doar în principiu, nu în timp real și, pe de altă parte, că metodologia logică este, în fond, incompletă.]**
- **Accentul pus o *metodologie inginerască* [care urmărește construirea de sisteme care funcționează în timp real].**

Sistemele expert, cazurile, rețelele de crezăminte, rețelele neuronale, programele evoluționiste, sistemele dinamice sunt toate ilustrări, într-un mod sau altul, ale obținerii rezultate efective, nu doar teoretice, chiar cu riscul de a jertfi ceva din universalitate.

Prezența acestui element de metodologie inginerască este, credem noi, firul care străbate tot arsenalul de unelte descris până acum. Pare cam ciudat să invoci metodologia inginerască atunci când este vorba de filosofie, dar tocmai aici constă noutatea.

În post-scriptum-ul la acest capitol care a căpătat deja dimensiuni cam mari vom trece în revistă critici aduse sistemelor artificiale și voi încerca să le contracarez prin punerea în evidență tocmai a rolului acestei metodologii inginerăști.

## Criticii sistemelor artificiale

Dintre criticii sistemelor artificiale voi reține trei nume proeminente: Hubert Dreyfus, John Searle și Roger Penrose. Cu argumente diferite, cei trei susțin că există activități intelectuale

propriu omului, activități care nu pot fi imitate de către sistemele artificiale.

Obiecțiile celor trei nu sunt interne, nu sunt obiecțiile cuiva care respinge un anumit mod de a construi sisteme artificiale pentru a propune altul. De pildă, Dreyfus se slujește de premise care-și au originea în opera lui Heidegger și Wittgenstein; concluziile sale sunt totuși diferite de cele ale susținătorilor sistemelor dinamice. Este vorba de un scepticism care vizează intenția de a construi sisteme inteligente ca atare.

Dreyfus și-a sintetizat critica sa la adresa "rațiunii artificiale", ca să folosim expresia sa, într-o carte celebră, apărută la începutul anilor 1970.<sup>128</sup> Din felul în care pune Dreyfus întrebările sale se vede imediat că atacul său se îndreaptă contra a ceea ce am numit aici inteligență artificială: (1) folosește omul, pentru a prelucra date, reguli formale, aidoma programelor de computer? (2) poate fi descris comportamentul omului cu ajutorul unui formalism, ce apoi ar putea fi transpus într-un program de calculator? Pentru a da substanță răspunsului său (negativ) la aceste întrebări, Dreyfus întreprinde o analiză a diverselor perioade de dezvoltare a inteligenței artificiale: (a) etapa modelării proceselor cognitive (1957-1962); (b) etapa simulării procesării informației semantice.

Dreyfus a pus în evidență contrastul dintre speranțele considerabile din prima perioadă și rezultatele modeste obținute (până în 1967). Astăzi multe dintre observațiile sale ar trebui cântărite cu grijă: programele profesionale de traducere fac, de pildă, o treabă foarte bună dacă textul sursă este scris clar (cu textele obscure au dificultăți și oamenii !); campionul mondial la șah a fost învins de un program de calculator, iar rețelele neuronale reușesc să fie poate mai utile decât un asistent uman managerilor. Au rămas însă și în 1999 multe dintre distanțele semnalate de Dreyfus între speranțele de început și rezultatele de pe parcurs.

<sup>128</sup> Hubert L. Dreyfus, *What Computers Can't Do. A Critique of Artificial Reason*, New York, Harper, 1972.

Partea mai profundă a investigației lui Dreyfus ține însă de analiza *presupozițiilor* inteligenței artificiale. Pentru ca sistemele artificiale să apară cât de cât umane, ele ar trebui să folosească limbajul. Dar, observă Dreyfus, pentru a atinge acest scop, cei din inteligența artificială încearcă să facă programul să urmeze reguli (formale). Or, susține Dreyfus, noi nu folosim reguli stricte atunci când utilizăm limbajul. Dreyfus îl citează pe Wittgenstein în sprijinul său.

Tot atât de greșită i se pare lui Dreyfus presupuziția conform căreia creierul uman ar funcționa ca un computer numeric. Funcționarea creierului ar semăna mai degrabă cu o combinație *sui generis* între un calculator numeric și unul analogic. Intuiția mea este că, după Dreyfus, ceea ce este foarte important e caracterul sui generis al creierului. Orice imitație ar trebui să urmeze îndeaproape originalul, pentru a obține rezultate similare. Situația aceasta ar face, evident, ca eventualele pretenții teoretice ale unei asemenea întreprinderi să fie foarte limitate. Dacă imit un tablou de Leonardo nu înseamnă că am o tehnică de a realiza în general tablouri ca Leonardo și nici măcar nu înseamnă că am o înțelegere a operei lui Leonardo. Falsurile sunt detectabile.

Dreyfus respinge și ceea ce el numește *presupoziția ontologică*, conform căreia lumea poate fi descrisă cu ajutorul faptelor atomare. Or, această presupuziție este acceptată atât în construcția părții fizice a sistemelor artificiale, cât și în partea lor impalpabilă (atomii de computații).

Omul, susține Dreyfus, pornește de la întreg pentru a înțelege detaliile. Sistemele artificiale pornesc de la detalii pentru a reconstitui întregul.

Omul se simte în lume ca acasă. El ființează *în* lume în sens heideggerian. Corpul omului nu poate fi separat de minte. Corpul contribuie la comportamentul uman. Fără mâini nu am fi ceea ce suntem.

La cea de a doua ediție a criticii sale la adresa rațiunii artificiale Dreyfus a adăugat o prefață,<sup>129</sup> în care analizează evoluțiile intervenite în inteligența artificială în anii 1970. Critica lui Dreyfus continuă să se concentreze asupra inteligenței artificiale bazată pe programe scrise în limbaje gen Lisp. El argumentează în continuare că și dacă am admite că rețeaua crezămintelor noastre poate fi formalizată complet, tot ar rămâne în afara formalizării contextul (dispozițiile noastre sufletești, grijile, abilitățile) fără de care crezămintele nu au înțeles.

John Searle, ca și Roger Penrose, au propus experimente gândite, menite să discrediteze inteligența artificială.

Experimentul lui Searle, numit “camera chinezească”,<sup>130</sup> constituie un contraargument la testul lui Turing.<sup>131</sup> Turing, creatorul *conceptului* de computer, și-a pus problema cum ar putea fi dat un sens operațional întrebării “Gândește computerul?”. Evident, el discuta lucrurile din perspectiva sistemului computațional ideal (denumit și “mașina Turing”),

<sup>129</sup> “From Micro-Worlds to Knowledge Representation: AI at an Impasse”, reprodusă în John Haugeland (ed.), *Mind Design*, MIT Press, 1981, pp. 161-204.

<sup>130</sup> Pentru o imagine la zi a dezbaterilor și bibliografiei consacrate experimentului imaginat de Searle vizitați pagina de web <http://www.uniroma3.it/kant/field/chinese.html>. Această pagină, semnată de Larry Hauser, este parte a ghidului de filosofia minții în construcție (în 1999) pe Internet la Universitatea Roma 3. Sursa întregii dezbateri o constituie articolul lui Searle “Minds, brains, and programs” *Behavioral and Brain Sciences* 3, 1980, pp. 417-424 [reprodus în Haugeland (ed.), *Mind Design*, pp. 282-306].

<sup>131</sup> Propus de Alan M. Turing în “Computing machinery and intelligence”, *Mind*, vol. LIX, nr. 236 (oct. 1950), pp. 433-460. O versiune în limba română, intitulată “Poate mașina să gândească?”, a apărut în volumul *Probleme filosofice și sociale ale ciberneticii* (seria “Dialectica marxistă și științele moderne”), vol. IV, București, Editura politică, 1963, pp. 305-330. Este interesant de amintit că acest volum venea ca o încercare de reparație a incredibilitelor condamnări ale ciberneticii în anii obscurantismului marxist. De aceea, retrospectiv, este o adevărată ironie că denumirea seriei conținea sintagma “dialectica marxistă”. Seria avea, de altfel, să capete, până la urmă, numele mai prozaic, dar onorabil, de “Știință și filosofie”.

dar avea speranța că, într-o zi, programe efective de computer vor putea trece testul următor:

A este un om, B un computer și C un om care nu știe dacă A sau B este computerul. C nu are posibilitatea de a intra în contact direct cu A sau B. El poate doar trimite bilețele cu întrebări (replici) lui A și B. Poate duce deci un dialog cu A și B, în urma căruia trebuie să decidă cine este omul și cine este computerul. Dacă mașina poate să-l păcălească pe C, atunci ea trece testul. Putem să ne imaginăm și grade de trecere a testului, în funcție de intervalul de timp în care mașina reușește să treacă drept om.

Pentru a înțelege argumentul lui Searle, să ne amintim puțin de operatorul uman care este capabil să introducă programul în memorie. Acest operator l-am înlocuit apoi cu sistemul de operare. Operatorul pune biții la locul lor, dar nu înțelege programul ca atare și nici nu are nevoie să-l înțeleagă pentru a-și desfășura activitatea.

În experimentul imaginar al lui Searle operatorul primește fișe pe care sunt scrise semne chinezești. Trebuie să transforme fiecare fișă într-un șir de biți (codul semnului chinezesc) și, în funcție de structura șirului de fișe primit, să introducă o serie de alte coduri (pentru structura gramaticală) și să adauge codurile dispozitivelor care prelucrează datele astfel obținute. Toate acestea sunt făcute de operator respectând regulile din "manualul operatorului", fără a pricepe nimic din înțelesul semnelor, deoarece – pentru operațiile pe care le face – operatorul nici nu are nevoie de așa ceva. De asemenea, dispozitivele sistemului evident nu înțeleg semnele și cu atât mai puțin șirurile de semne. Putem, de altfel, presupune că șirul complet nici nu mai trece printr-un singur dispozitiv. Ceea ce produce sistemul este un nou șir de coduri, pe care operatorul le transformă acum în fișe pe care sunt scrise cuvinte în limba engleză (sau română – operatorului îi este indiferent care este limba la ieșire!). Aceste șiruri de fișe sunt transmise în exterior ca rezultat.

Cum rezultatul ar putea fi nu doar o traducere, ci și o repovestire, o rezumare, o explicare etc., deci o probă de *înțelegere* a textului chinezesc, Searle trage concluzia că sistemul artificial ar trece testul lui Turing fără să fie o minte.

Searle nu respinge ideea că programele de calculator ar putea fi unelte pe care noi le putem folosi pentru a înțelege mintea umană, dar respinge ipoteza că ar putea fi produse minți artificiale.

Roger Penrose a analizat un experiment gândit menit să ne arate ce ne-ar împiedica să construim un duplicat al unei minți umane date.<sup>132</sup> Experimentul constă în *rasterizarea* (în sensul metaforic folosit aici, cel de transformare în șiruri de biți) proprietăților atomilor creierului cuiva și transmiterea la distanță a șirurilor de biți, unde totul va fi recompus așa cum a fost. Putem, desigur, să ne închipuim că acest lucru a fost făcut cu tot corpul persoanei respective și că se urmărește reconstituirea exactă a configurației fizice în alt loc.

Testul lui Turing ar putea fi trecut de duplicat, ca și cum ar fi originalul, dar aceasta nu arată decât – crede Penrose – caracterul deficitar al testului, care este un test pur operațional. Copia nu este originalul.

În al doilea rând, apar probleme legate de *identitatea personală*. Ce ar zice originalul dacă i s-ar spune că a devenit de prisos din moment ce există o copie la fel de bună ca originalul? Originalul are, metafizic, prioritate și copia, din nou, cum fi la fel de bună cu originalul?

Dar cheia argumentului teleportării, după Penrose, pare să fie legată de fenomene fizice particulare proprii doar unui anume creier. Aceste fenomene sunt de natură cuantică și există o problemă când este vorba de copierea lor.<sup>133</sup> Nu putem face o copie a unei stări cuantice fără ca să o modificăm. Problema identității nu s-ar mai pune, deoarece nu există decât posibilitatea de a avea un singur exemplar. Duplicarea este exclusă.

---

<sup>132</sup> Roger Penrose, *Mintea noastră ...cea de toate zilele*, București, Editura Tehnică, 1996, pp.35-37. Titlul în original al cărții lui Penrose este *The Emperor's New Mind*.

<sup>133</sup> Penrose, *Mintea noastră ...cea de toate zilele*, pp.293-294.

Penrose pare să ezite să propună drept argument natura foarte particulară a fenomenelor cuantice din creier.<sup>134</sup> El se concentrează asupra a ceea ce el socotește că este natura *nealgoritmică* a gândirii noastre conștiente. Pentru aceasta el recurge la faimoasele teoreme ale lui Gödel, la care a revenit într-o carte mai nouă.<sup>135</sup>

Ar însemna să facem un ocol cam mare aici dacă am încerca să explicăm teoremele lui Gödel pe care se sprijină argumentarea lui Penrose. Cea mai cuminte presupunere este aceea că cititoarea sau cititorul cunosc aceste teoreme din logică. Oricum, ceea ce vrea să spună Penrose este relativ simplu: conștiința permite acțiuni care trec dincolo de orice activitate computațională. Penrose susține că argumentele sale extind argumente anterioare formulate de Gödel însuși, de Nagel și Newman, precum și de J. R. Lucas, profesor și el (de filosofie) la Oxford, ca și Penrose (care este profesor de matematică).

În esență, argumentul lui Penrose poate fi schițat după cum urmează. Știm cu toții un anume gen de învățare a matematicii: studiezi bine tipuri de probleme și afli care este metoda care rezolvă problemele de tipul respectiv. Dacă primești spre rezolvare o problemă, atunci cauți să vezi de ce tip este și astfel descoperi prin ce metodă o poți rezolva. Să zicem că sistemul SMM încapsulează totalitatea metodelor matematice cunoscute. Mai precis, extindem la limită această mulțime. Penrose presupune că SMM este un sistem formal (în sens logic). Pare a avea însă noimă să considerăm că SMM este un sistem computațional în a cărui memorie au fost introduse ca programe toate metodele matematice (introduse într-un mod nu neapărat computațional, precizează Penrose). Fie acum un matematician ipotetic M. M nu știe în mod necesar toate metodele matematice. Dar putem gândi că modelul cognitivist al minții ar putea să confere un sens cât se poate de exact

<sup>134</sup> Creierul este "prea cald" pentru a menține coerența cuantică (Penrose, *Mintea noastră ...*, p.434).

<sup>135</sup> Roger Penrose, *Shadows of the Mind. A Search for the Missing Science of Consciousness*, Oxford University Press, 1994. Cartea este disponibilă și în traducerea românească a Danei Jalobeanu.

propoziției “Eu sunt SMM”. Un matematician imaginar ar putea aserta această propoziție. Dacă o face, el extinde însă SMM la sistemul SMMC (SMM plus aserțiunea – adevărată – că este SMM). Sistemul SMM trebuie să fie consistent (în sensul că nu generează aserțiuni contradictorii – dacă ar fi inconsistente ar produce orice fel de aserțiuni și nu ar prezenta deci interes). Dacă SMM este consistent, atunci și SMMC ar trebui să fie consistent. Să spunem că există o “rasterizare” (o codificare) a arhitecturii SMM (inclusiv conținutul memoriei SMM). Putem folosi acest cod pentru a calcula rezultate privitoare la SMM pe o mașină virtuală  $G(SMM)$ . Între altele, se poate stabili consistența lui SMM. “G” vine de la numele lui Gödel. Principalul rezultat în privința  $G(SMM)$  este că mașina virtuală  $G(SMM)$  nu poate fi construită pe arhitectura SMM. Același lucru îl putem spune și despre  $G(SMMC)$ . Dar unul dintre rezultatele  $G(SMMC)$  este “Eu sunt SMM”, pe care matematicianul M nu-l poate deci computa. Dar M are ideea respectivă. Concluzia lui Penrose este că M folosește metode necomputaționale pentru a ajunge la ideea respectivă.

Cartea lui Penrose a fost subiectul unei substanțiale polemici în revista *Psyche*.<sup>136</sup> Penrose a răspuns unor critici venite din cele mai diferite direcții. El a recunoscut o serie de defecte de ordin tehnic,<sup>137</sup> dar în rest și-a dezvoltat și apărut poziția din carte.

Teoria conștiinței, după Penrose, trebuie să facă apel la fizică,\* mai precis la mecanica cuantică. Funcționarea la nivel cuantic a sinapselor dintre neuroni ar explica aspectele non-computaționale ale creierului.

---

<sup>136</sup> *Psyche* ianuarie 1996. Vezi pe web <http://psyche.cs.monash.edu.au/v2/>. Vezi fișierul [psyche-2-23-penrose.html](http://psyche.cs.monash.edu.au/v2/psyche-2-23-penrose.html) pentru răspunsurile lui Penrose la critici [data accesării de către noi 27.04.1998].

<sup>137</sup> Semnalate în special de S. Feferman. Cea mai importantă eroare privește conceptul de omega-consistență, folosit de Penrose în carte. Dacă expunerea noastră de mai sus ar fost tehnică și nu intuitivă, ar fi trebuit desigur ținut cont de aceste observații. Oricine studiază independent teoremele lui Gödel va da oricum peste acest concept.



## Și totuși atracția abordărilor compuționale

Nici un critic nu contestă valoarea computerelor ca unelte de studiere a minții. Ar fi și fără sens să scrii așa ceva. Cercetările experimentale în psihologie au enorm de câștigat, ca peste tot în știință, de pe urma computerelor. Simulările pe computer sunt extrem de prețioase.

Ceea ce resping criticii este ideea că simularea pe computer a comportamentelor umane ar însemna și posibilitatea de a crea *minți artificiale*. Dacă simulez un vulcan cu ajutorul unui program nu înseamnă că erupe muntele în calculator.

Am putea pune punct aici discuției și accepta observația de mai sus ca pe una de bun simț, dacă nu ar fi și în opera criticilor o serie de presupoziii discutabile.

Dreyfus, între altele, susține că poziția sa este antimecanicistă. Presupozitia pe care o văd eu aici în spatele acestei susțineri este că nu există nici o deosebire între ceasul mecanic și computer. Nu vreau să apăr acum meritele ceasului mecanic, o construcție remarcabilă și cu influență probabil și asupra filosofiei politice (reglarea funcționării statului prin balansarea și frânarea puterilor). Ceea ce scapă în punerea tuturor mecanismelor în același coș este caracterul universal al computerului. Prin programare, această mașină poate fi transformată în orice altă mașină.

Searle are, în parte, dreptate când critică testul lui Turing. Dar ceea ce pune el în evidență este o presupozitie behavioristă a testului. Cognitivismul înseamnă, la urma urmei, că mentalul contează. Nu contează doar comportamentul. În sensul acesta, testul lui Turing nu este congruent cu articulațiile mai profunde ale concepției acestuia.

Demonstrația lui Penrose, inspirată de teoremele lui Gödel, este corectă. Imperfecțiunile ei de ordin tehnic nu cred că ar schimba prea mult tăria concluziei. Presupozitia însă admisă tacit este că un sistem compuțional nu este *nimic altceva* decât

un sistem logic. Metodologiei ingineresti îi este acordat prea puțin credit.

Oricine a scris vreodată un program știe că familiarizarea cu logica este avantajoasă, dar nu suficientă. Fără metodologia inginerască adecvată, programul nu poate fi construit. De aceea este cel puțin ciudat că idei precum abstracția sau modularitatea, în sensul metodologiei ingineriei programelor nu-și au locul în discuțiile critice.

La urmă, dar nu în cele din urmă, propunerea lui Penrose de a recurge în mod decisiv la mecanica cuantică pentru a explica mintea are drept presuposiție credința că rafinamentele materiei ar fi de preferat rafinamentelor ideilor.

Dacă ținem cont de aceste contraobservații critice, cred că putem merge cu ceva mai multă speranță mai departe. Cum lecția criticilor trebuie totuși luată în serios, am separat în cele ce urmează zonele în care există rezultate palpabile, de cele în care cercetarea minții este confruntată cu mistere ce par încă impenetrabile.

## Abordări computaționale ale minții

Cititorii cu interese filosofice care au sărit peste cel de-al treilea capitol ar trebui avertizați că au pierdut nu doar ocazia de a-și spori bagajul de unelte conceptuale de care s-ar putea să aibă acută nevoie în continuare, ci și rădăcinile discuției filosofice.

Filosofia minții, privită ca o componentă a științei cogniției, întreține cu arta construirii de sisteme artificiale inteligente o relație aparte. Asupra ei a atras de multă vreme atenția John McCarthy, creatorul Lisp-ului. Iată ce scria McCarthy:

*Filosofia are cu inteligența artificială o relație mult mai directă decât cu alte științe. Ambele discipline presupun formalizarea cunoașterii comune și remedierea defectelor acesteia. Deoarece un robot dotat cu o capacitate generală de a fi inteligent presupune inculcarea unei anumite viziuni despre lume, deficiențele în investigarea introspectivă de către programator a propriei sale lumi se vor traduce într-o slăbiciune operațională a programului.*<sup>138</sup>

Se impun aici mai multe remarci. Prima este că putem reține fără rezerve observația inițială a lui McCarthy. Semnificația ei este următoarea: filosofia întreține cu științele naturii o relație mai degrabă indirectă (ea analizează structura teoriilor acestora după ce ele s-au constituit sau contribuie la formarea presupuzițiilor cunoașterii în aceste domenii, dar nu direct la construirea teoriilor); în cazul științelor sociale relația este mai complicată și filosofia pare implicată mai adânc, dar nu în modul în care acest lucru se petrece în inteligența artificială. După cum atrage atenția McCarthy, programele de computer

---

<sup>138</sup> John McCarthy, "Philosophical Notes", in *Epistemological Problems of Artificial Intelligence* 1977, <http://www-formal.stanford.edu/pub/jmc> [transpus în HTML de John McCarthy la 15 mai 1996].

sunt deficitare dacă filosofia este ignorată sau făcută prost. Motivul este aproape evident. Este vorba în primul rând de faptul că aceste programe implică operarea într-un fel sau altul cu *cunoaștere*. De asemenea, viziunea asupra lumii – după cum precizează McCarthy – are un rol important.

Cei care au parcurs capitolul anterior pot aprecia, pe de altă parte, cât de frământată este istoria construcției de sisteme inteligente și cât de multe sunt ecourile filosofice ale evoluțiilor din acest domeniu.

Ecourile în filosofia minții sunt de-a dreptul ininteligibile fără o familiarizare cu sistemele inteligente. Cel mai bun exemplu este chiar termenul de *computaționalism* (atenție, în această carte se face o distincție între “computațional” și “computaționalist” !). Definițiile sunt inutile pentru a înțelege ceva. Ele sunt simple rezumate a ceea ce deja ai priceput. Iar, în acest caz, o minimă familiarizare cu programarea în Lisp este absolut necesară. Așa ai înțelege mai bine de ce computaționalismul este denumit uneori și *symbolicism*. Încercarea de a înțelege toate acestea plecând doar de pe platforma logicii simbolice nu mi se pare inspirată. Cunoașterea logicii simbolice este utilă, dar nu suficientă, din motive explicate deja (nevoia de a înțelege metodologia inginerască).

Rezultatele obținute sunt fragmentare. Unele zone ale minții sunt iluminate de teorii care, fără să fie infailibile, sunt testabile cu mijloacele uzuale în comunitățile științifice. De asemenea, o serie de dispute filosofice au prins un contur limpede și taberele par să fi intrat în faza unui “război de poziții” (fiecare și-a săpat tranșeele și încearcă de acolo să bombardeze din când în când tabăra adversă). Printre progresele clasice se situează rezultatele lui Marr în teoria vederii, dar și o serie de cercetări referitoare la memorie și, evident, limbaj. Războiul filosofic de poziții cel mai aprig pare să se poarte între computaționaliști și conexioniști, dar și alte subiecte cum ar fi ipoteza limbajului gândirii formulată de J. Fodor sau modularitatea minții sunt viu dezbătute.

Dintre zonele misterioase ale minții iese în evidență, în primul rând, conștiința, dar și teme cum ar fi "gândirea" sau limbajul intern.

Capitolul de față se concentrează asupra a ceea ce poate fi spus cu mai mare limpezime despre minte și cred că acest supliment de claritate este meritul științei cogniției și rodul folosirii noilor unelte conceptuale.

### **Vederea ca un proces computațional**

Mulți oameni, când privesc un film științifico-fantastic în care computere superdotate văd tot ce se petrece în jur, nu par a fi surprinși. Deși cei mai mulți dintre ei ar respinge perspectiva unui computer care "se comportă uman", care - de pildă - are conștiință, ei nu se revoltă la ideea că un computer ar putea vedea.

A vedea este departe de a fi însă un lucru foarte simplu. Cum de nu avem în fața ochilor doar pete de culoare amorfe? Cum de vedem obiectele în spațiu? Sau, mai complicat, ne slujim de vedere pentru a ne feri de pericole sau pentru a afla ce este în lume? Ce ar însemna frumusețea unei imagini? Înțelegerea capacității de a vedea presupune o mai profundă pătrundere a mecanismelor minții decât ni se pare, poate, la o primă examinare a lucrurilor.

La aceste întrebări răspunsurile au venit din multe direcții: neuroștiință, psihologie, estetică. Aceste paragrafe propun, mergând pe urmele deja clasicei contribuții a lui David Marr, un răspuns complex, structurat din perspectiva științei cogniției. Acest gen de răspuns rezumă bine felul în care noua abordare, țesută pe urzeala oferită de înțelegerea proceselor computaționale, a integrat contribuțiile mai multor discipline.

Marr, care a trăit între 1945-1980, a pus o piatră de hotar, în anii șaptezeci ai secolului al XX-lea, la înțelegerea de

către noi a minții.<sup>139</sup> Ideile sale vor ilustra felul în care unele conceptuale prezentate în capitolul al treilea al cărții servesc la constituirea perspectivei cognitiviste asupra minții.

Marr pune în contrast interesele de cunoaștere ale cercetătorilor sistemului nervos cu cele ale celor preocupați de procesele computaționale. Pentru primii este foarte important să afle cât mai multe despre celulele care alcătuiesc sistemul nervos, felul în care sunt conectate, reacțiile lor la diverse stimulări. Pentru ceilalți, întrebarea ar fi mai degrabă una cu privire la felul în care stimulările sunt tratate de către sistem ca date și supuse unui proces de prelucrare. Dacă vederea este prelucrare de date, atunci ar fi posibil să conectezi cumva o cameră de luat vederi cu un computer și să programezi vederea.<sup>140</sup>

Explicația urmează să vină, pentru David Marr, din multiple direcții. Fiecare dintre părțile menționate mai sus are contribuția sa. Rezolvarea de către creier a unei sarcini legate de prelucrarea de informații are nevoie de diferite genuri de explicații.<sup>141</sup>

Adaptând ideile lui Marr,<sup>142</sup> putem spune că a înțelege vederea ca proces computațional înseamnă: (1) să identificăm problema(ele) vederii; (2) să delimităm condițiile pe care trebuie să le îndeplinească soluția (care sunt datele de intrare și cele de ieșire, care sunt țelurile procesării și strategia procesării datelor);

---

<sup>139</sup> David Marr și-a prezentat concepția într-o formă accesibilă, în cartea sa *Vision - A Computational Investigation into the Human Representation and Processing of Visual Information*, New York, W.H. Freeman and Company, 1982. Prezentări scurte, dar mai dezvoltate decât cea oferită aici, ale concepției lui Marr se găsesc în Philip Johnson-Laird, *The Computer and the Mind*, ed. a II-a, Londra, Fontana Press, 1993, pp.57-126 și, în limba română, în cartea lui Mircea Micle, *Psihologie cognitivă*, Cluj-Napoca, Gloria, 1994, pp.76 ș.u.

<sup>140</sup> Vezi David Marr, *Vision*, p.4.

<sup>141</sup> Cf. David Marr, *Vision*, pp.4-5.

<sup>142</sup> Marr insistă asupra ideii de proces. Ducerea la bun sfârșit a unei sarcini de procesare a datelor are, la el, trei niveluri, botezate "computațional", "reprezentare a datelor de intrare și de ieșire plus algoritmul de prelucrare a lor", "implementare la nivel hard". V. David Marr, *Vision*, pp.22-25.

(3) să precizăm modul în care sunt prelucrate datele; (4) să găsim felul în care procesul computațional poate fi realizat fizic. Toți acești pași amintesc foarte bine de individul care calcula noul preț al pâinii în condițiile inflației: (a) acesta avea problema stabilirii noului preț al pâinii; (b) datele sale de intrare erau prețul vechi și sporul pricinuit de inflație, iar datele de ieșire trebuie să-i furnizeze prețul nou; (c) prelucrarea însemna în esență o adunare; (d) totul se putea face de către un om înarmat cu niște fișe și un creion.

### Filtrarea critică a ipotezelor cu privire la vedere

De multe ori, când le vorbești oamenilor despre abordarea cognitivistă a minții în general și a vederii în particular, reacția lor este să se întrebe dacă toate aceste încercări au vreo relevanță pentru ceea ce se petrece "în minte". Filosofic, întrebarea aceasta este interesantă. Ce legătură este între computații pe calculator și vederea reală? În ce sens preluăm noi date din mediu și le prelucrăm? Se petrec oare lucrurile în minte așa cum sunt zugrăvite ele în modelele cognitiviste (bazate pe ideea proceselor de calcul)?

Versiunea noastră a răspunsului la aceste întrebări este că noi ne slujim de procesele de calcul imaginate de către noi așa cum ne slujim de orice încercare de a ghici cum stau lucrurile, de orice ipoteză într-un experiment.

Experimentele au o funcție critică. Ele selectează tentativele noastre de a prezice/explica fenomene. Experimentele nu ne spun deci cum stau lucrurile, ci cum nu stau lucrurile. Toată strădania este să filtrăm treptat ipotezele noastre (cu privire la vedere, de pildă). Noi serii de experimente supun unei critici suplimentare severe ipotezele formulate. Aceste teste sunt sugerate de teorii diverse despre minte, cum ar fi cele ale neuroștiinței sau ale psihologiei ș.a.m.d.

Un exemplu simplu de filtrare a unei ipoteze ne este oferit de către Philip Johnson-Laird. Vederea ar însemna în

perspectiva ipotezei examinate, producerea de imagini, cam așa cum producem tablouri pe care le atârnăm într-o galerie. Filtrul despre care este vorba în acest caz este cel al unei critici filosofice: dacă imaginile sunt atârnate în galeria minții, atunci cine și, mai ales, cum le vede? Explicația văzului nu ar fi decât amânată!<sup>143</sup>

Mai mult decât atât, procesele de calcul imaginate pot fi simulate pe computere. Putem astfel să ne "jucăm" în diverse moduri cu ele. Important este, până la urmă, dacă ipotezele noastre trec de filtre de teste cât mai severe.<sup>144</sup>

### Problemele vederii

Vederea, scrie Marr, poate fi folosită pentru a soluționa cele mai variate probleme: porumbeii se slujesc de ea pentru a naviga în aer, un păianjen are un mecanism special pentru a distinge între un partener la împerechere și o posibilă pradă, iepurele se pare că are pe retină detectori speciali care recunosc forma unui uliu.<sup>145</sup> Omul are un văz mai puțin specializat, deși anumite mecanisme speciale sunt prezente (de exemplu, când închidem ochiul la apropierea prea rapidă de cap a unui obiect).

Toate acestea sunt probleme la a căror rezolvare contribuie vederea. Dar, în același timp, ele iluminează felul în care se constituie problematica vederii ca atare. Astfel, ca să mă feresc de un obiect ce vine rapid către mine, trebuie să procesez anumite date ce se găsesc în mediul ambiant. Razele de lumină reflectate de obiect sunt astfel de date.

<sup>143</sup> Cf. Philip Johnson-Laird, *The Computer and the Mind*, p.57.

<sup>144</sup> Karl Popper a dezvoltat cel mai bine acest punct de vedere al "învățării din greșeli". Schema popperiană este aici adaptată, căci avem de a face cu mai multe teorii și cu efortul de a le testa critic. Fiecare dintre teoriile noastre, luată separat, este însă greu - dacă nu imposibil - de supus unor teste critice. Sub foc critic încrucișat, ideile noastre pot fi trecute prin ciurul discuției critice și avem astfel calea deschisă către cunoaștere.

<sup>145</sup> David Marr, *Vision*, p.32.



Voi prelucra oare raze de lumină ca atare? Nu. Atât omul, cât și roboții, de pildă, transformă energia luminoasă în semnale electrice.<sup>146</sup> Semnalelor acestora (datele prelucrabile de robot sau mintea omului) trebuie să fie însă cumva distinse unele de altele, grupate, trebuie găsite forme în lumea din jur ș.a.m.d. Acestea sunt problemele vederii.

În rezumat, în primă instanță, problema vederii este să procesezi un set de semnale electrice. Este aici o formă specifică a "rasterizării" prezentate până acum doar în principiu. Marr descrie aceste semnale primare înstructurate ca într-un tabel. O poziție  $(x,y)$  din tabel va însemna că ne aflăm în punctul sau celula de la intersecția rândului  $x$  cu coloana  $y$ . Intensitatea seninalului electric în acest punct este notată cu  $I(x,y)$  și este chiar valoarea înscrisă în punctul  $(x,y)$  în tabel.<sup>147</sup>

Din punctul de vedere al procesului computațional, lucrurile apar ca și cum ne-am afla în fața unui imens tabel cu numere reprezentând fiecare intensități (ale luminii, în ultimă instanță).

Problema noastră imediat următoare este să localizăm schimbări de intensitate.

### Schița primară: pierdere de date și câștig de structurare

Schimbările de intensitate din tabelul descris mai sus sunt însă adesea prea puțin nete de la o celulă la alta și au toate șansele să fie încețoșate de "zgomote". Ideea lui Marr este că procesul de calcul transformă aceste treceri graduale în contraste. Imaginea continuă este discretizată.<sup>148</sup>

---

<sup>146</sup> V., pentru mai multe detalii, Philip Johnson-Laird, *The Computer and the Mind*, p.59 ș.u.

<sup>147</sup> David Marr, *Vision*, p.31.

<sup>148</sup> V. David Marr, *Vision*, în special p.67 (pentru modul de a rezolva problema), precum și paginile precedente pentru detaliile tehnice.

Structura de tip tabel (cu intensități) capătă o formă mai complexă. Elementele discrete care rezultă au un set de descriptori care sunt aidoma celor de care ne-am putea sluji pentru a trasa o linie pe foaia de hârtie sau pe ecranul computerului. Am putea indica, de pildă, următoarele:

ORIGINE \_\_\_\_\_

DIRECȚIE \_\_\_\_\_

LUNGIME \_\_\_\_\_ .

Cuplând elemente discrete, putem constitui forme mai complicate.

Interesant este că, deși am pierdut date (deoarece, prin discretizare, s-a pierdut continuumul inițial, constând - la origine - din raze de lumină), se câștigă ceva în planul constituirii imaginii. După Marr, noi nu putem sesiza nemijlocit discretizarea primară, ci doar descrierea care ia naștere în ceea ce el a denumit schiza primară.<sup>149</sup>

## Modularitatea văzului

Schița primară este aidoma unui tipar al contururilor. Ea seamănă oarecum cu desenul pe care un pictor îl face pentru a realiza apoi un tablou. Peste schița făcută sunt așezate culorile, este dozată luminozitatea ș.a.m.d.

Vederea umană este capabilă să sesizeze adâncimea, mișcarea, forme și umbre, textură și culori. După Marr, fiecare dintre aceste capacități este realizată de către un modul. Modulele acestea funcționează în paralel.

Vederea în adâncime Marr o denumeste, cu un termen compus din elemente grecești, stereopsis. Computații specifice

---

<sup>149</sup> Cf. David Marr, *Vision*, p.74. Marr arată cum imaginea discretizată (practic, un raster cu puncte foarte mari) a lui Abraham Lincoln, produsă de L.D. Harmon, nu poate fi văzută (de aproape). Dacă îndepărtăm imaginea (și pierdem date, căci nu mai distingem exact fiecare pătrățel din care ea este compusă), atunci sesizăm figura lui Lincoln.

sunt presupuse pentru a da seamă de capacitatea noastră de a sesiza adâncimea. De pildă, capacitatea noastră de a sesiza profunzimea este pusă de cele mai multe ori pe seama faptului că avem doi ochi care produc imagini ușor diferite (dacă ochii ar fi dispuși lateral, imaginile s-ar suprapune relativ puțin).

Marr pornește și el în considerațiile sale de la observații de genul celor de mai sus. Ipoteza este că disparitatea creată între imaginile de la cei doi ochi este exploatată în computații pentru a realiza vederea în adâncime (pentru a obține, cu alte cuvinte, date cu privire la dispunerea în spațiu a obiectelor).

Modulele celelalte funcționează și ele, ca să spunem așa, tot exploatând "indicii", "fire" oferite de datele de care dispunem după construirea schiței primare.

Prin urmare, analogia cu tabloul pictorului are limitele ei, dacă ne-am gândi că schița primară este umplută (oarecum liber) ca și în cazul tabloului. Văzul, în felul în care îl teoretizează Marr, seamănă cu felul de a proceda al unui pictor realist. Obiectele așa cum sunt ele îl constrâng. Sunt exploatate toate datele care sunt deja disponibile, pentru a completa imaginea. Totul este aidoma unei operațiuni arheologice.

De ce aș vedea, de exemplu, o suprafață ondulată? Contururile sinusoidale sesizate de către noi servesc drept indicii care conduc la perceperea unor undulații. Teoria rezistă și la experimente care ar putea-o infirma: contururi sinusoidale trasate pe hârtie ne fac să vedem o undulație chiar și acolo unde ea nu există.<sup>150</sup>

Ce anume face ca un om sau un cal să se detașeze de fondul imaginii? De ce, la urma urmei, să vedem obiecte distincte unele de altele? Marr furnizează argumente menite să ne convingă că separarea obiectelor ca atare unele de altele, într-o scenă, nu este realizabilă doar pe baza datelor cu privire la diferențele de intensitate a luminii.<sup>151</sup>

---

<sup>150</sup> Cf. David Marr, *Vision*, p.227.

<sup>151</sup> Pentru explicarea problemei v. David Marr, *Vision*, p.272-273.

Modulele exploatează, cum arătam mai sus, diferite indicii, diferite "fire" pe care le oferă datele de intrare. Cum pot fi însă coordonate toate rezultatele obținute?<sup>152</sup>

Marr a formulat ipoteza că procesul de integrare este în două trepte. O primă fază constă din inferențe cu privire la continuități și discontinuități în spațiu.<sup>153</sup> În felul acesta se articulează reprezentări ale unor obiecte distincte.

Faza a doua înseamnă un nivel de integrare la care suntem capabili să recunoaștem obiecte. Aceasta este de altfel ținta finală a investigației lui Marr.

Pentru a recunoaște obiecte nu ajunge să înțelegem o imagine, ca în procesul de separare a obiectelor de fundalul unei scene. Marr presupune că avem și un catalog de modele, organizate modular, care ne permit să efectuăm recunoașterile.<sup>154</sup> O girafă din catalogul lui Marr, de pildă, seamănă cu orice alt patruped, atâta doar că are gâtul sensibil mai lung.

### În loc de rezumat: rolul cunoașterii în cadrul vederii

Philip Johnson-Laird atrage atenția asupra faptului că problema vederii este insolubilă, dar rezolvată totuși de organisme care văd.<sup>155</sup>

Problema vederii este insolubilă pentru că ea înseamnă reconstituirea, pornind de la înstructurări de raze de lumină, a obiectelor care au generat înstructurările respective. Or, mai multe obiecte ar fi putut genera înstructurarea dată.

Ea este rezolvată totuși de către organisme care văd, crede Johnson-Laird, pentru că acestea recurg la cunoaștere. Cunoașterea de care vorbește Johnson-Laird o putem

<sup>152</sup> Cf. David Marr, *Vision*, p.276.

<sup>153</sup> V. cazul perechii stereo a lui Andrew Witkin, în care imaginea (bidimensională) ne permite să inferăm existența a două benzi suprapuse (cf. David Marr, *Vision*, p.287).

<sup>154</sup> David Marr, *Vision*, p.305 ș.u.

<sup>155</sup> Philip Johnson-Laird, *The Computer and the Mind*, p.125.

interpreta ca pe o capacitate de a folosi și genera procese computaționale.

## Structurile de date și memoria

Acest capitol ilustrează, poate cel mai insistent, intenția de bază a cărții: aceea de a pune în mâna cititoarei/cititorului un set de unelte și de a ilustra apoi felul în care acestea au fost folosite în cercetarea minții. Identificarea structurilor de date reprezintă unealta aflată în paragrafele imediat următoare în centrul atenției. Prin urmare, nu descrierea exhaustivă a tipurilor de memorie, ci prezentarea felului în care structurile acestea constrâng procese mentale reprezintă aici principalul obiectiv.

În cazul vederii, am constatat cum, la un moment dat, pierderile de date sunt contrabalansate printr-un câștig în planul structurării. *Mutatis mutandis*, există un fenomen familiar și în cazul memoriei. Uităm parcă nu atât pentru a face loc unor date noi, ci pentru că memoria umană este un proces activ de înstructurare a datelor. Nu ținem minte totul și nu aruncăm datele în memorie ca într-o magazie de amintirile care stau claie peste grămadă.

Dispozitivele de memorare, pe de altă parte, sunt dispozitive esențiale în cazul unui proces computațional. Iar dacă urmărim analogia cu computațiile mai îndeaproape, memoria servește atât la stocarea datelor, cât și a programelor de prelucrare a lor.

Sisteme de dispozitive de memorare variate par a se afla și în arhitectura minții umane. Câte numere puteți aduce de-o dată în "fața ochilor minții"? Răspunsul este: șapte, plus sau minus două numere.<sup>156</sup> Nu prea multe! Puteți face același experiment cu altceva decât numere: fețele oamenilor, coperte de cărți. Încercați, de asemenea, să priviți numere și să le

---

<sup>156</sup> Este vorba de răspunsul cuprins în studiul lui George A. Miller "The magical number seven, plus or minus two", publicat în 1956 în *Psychological Review* (apud Philip Johnson-Laird, *The Computer and the Mind*, p.147).

repetată (fără să vă uitați la ele). Din nou, numărul magic șapte își face apariția.

Memoria la care facem apel în exercițiile de mai sus este memoria de scurtă durată, fără structuri de date bogate. În schimb, memoria de lungă durată este extrem de cuprinzătoare. S-ar putea să nu ținem minte multe numere de telefon, date istorice sau date de naștere, dar tot știm un număr considerabil (și astfel depășim acel șapte fatidic al lui George Miller). Cu toții ținem minte o mulțime de lucruri. Evident, nu le putem aduce însă simultan în memoria de scurtă durată.

Memoria de scurtă durată a fost, la rândul ei, subdivizată în mai multe tipuri. Avem, de pildă, nevoie de o memorie senzorială de scurtă durată pentru a reține schița primară a unei imagini vizuale. Memoria aceasta păstrează schița până când procesarea datelor senzoriale venite de la retină atinge următorul stadiu.<sup>157</sup> Fără acest tip de memorie, am ajunge în următorul stadiu al procesului vederii să sesizăm doar pete de culoare ș.a.m.d.

Până acum am vorbit despre dispozitivele de memorare din arhitectura minții. Să examinăm însă și structurile datelor stocate de aceste dispozitive. După cum am văzut, structurile de date sunt o parte a unui program. Alegerea unor structuri potrivite face ca și algoritmul de prelucrare a datelor să funcționeze eficient. În genere, ca parte a unui program, structurile de date constrâng procesul computațional, fac ca nu toate opțiunile posibile să fie încercate. Dacă asemenea structuri există în memoria umană, atunci ele ar trebui să genereze același efect de constrângere, care poate fi pus în evidență experimental.

Structura de date examinată la care ne vom referi în continuare este stiva (v. despre ea mai multe în capitolul al treilea).

Ideea fundamentală este că o structură de tip stivă care ar fi prezentă în memoria umană ar constrânge posibilitățile noastre de a construi și înțelege propoziții declarative, întrebări

---

<sup>157</sup> Cf. Philip Johnson-Laird, *The Computer and the Mind*, p.149.

etc. Dacă elementele propoziției sau întrebării trebuie ținute minte pentru scurt timp, deoarece sunt necesare în procesul de înțelegere a propoziției, atunci contează foarte cum sunt structurate (dispuse într-o structură de date) în memorie aceste elemente.

### Experimentele lui Wanner și Marantz: punerea în evidență a structurilor tip stivă

Wanner și Marantz au realizat experimente în care modul de a construi propoziții este explicat cu ajutorul unei structuri a memoriei pe care ei au denumit-o HOLD (reține!) și care este necesară pentru a păstra elementele propoziției în memorie pe parcursul procesului de a analiza a înțelesului acesteia.<sup>158</sup>

O consecință a ipotezei existenței memoriei HOLD este că întreruperea enunțării unei propoziții în momentul în care procesul de detectare a unei părți a înțelesului nu s-a încheiat va solicita mari eforturi din partea subiecților, pentru că memoria HOLD este puternic solicitată (probabil dincolo de capacitățile și destinația ei de bază). Faceți o comparație în acest sens între următoarele propoziții din limba engleză, a căror enunțare este întreruptă de enumerarea unei liste de nume în punctele marcate de paranteza care conține o steluță:

(1) *The witch who despised (\*) sorcerers frightened little children.*

(2) *The witch who sorcerers (\*) despised - frightened little children.*

---

<sup>158</sup> E. Wanner și M. Marantz, "An ATN approach to comprehension", în M. Halle, J. Bresnan, G. Miller (eds.), *Linguistic Theory and Psychological Reality*, Cambridge (SUA), MIT Press, 1978. Rezultatele lui Wanner și Marantz sunt descrise în Edward Matthei și Thomas Roeper, *Introduction à la psycholinguistique*, Paris, Dunod, 1988. Am extras exemplele folosite în text din cartea lui Matthei și Roeper.

Înțelesul enunțului (1) este, cu aproximație: vrăjitoarea care-i disprețuia pe vrăjitori speria copiii mici. În timpul experimentului, subiecților, după secvența "the witch who despised", li se dă o listă de cinci nume proprii. Apoi este reluat enunțul (1).

Înțelesul enunțului (2) este: vrăjitoarea pe care vrăjitorii o disprețuiau speria copiii mici. Problema cu care se confruntă subiecții în al doilea caz este că trebuie să facă apel la memoria HOLD. "The which who despised" este o sintagmă care are deja un înțeles conturat (este o "vrăjitoare care disprețuiește"). Dar "the which who sorcerers" nu te lasă să întrevezi despre ce este vorba și elementele trebuie ținute în memoria HOLD până când noi elemente vor putea contura un înțeles. Efortul cerut de memoria HOLD face ca să fie mai greu de reținut lista de nume proprii.

Cum ar putea fi pusă însă în evidență structura memoriei HOLD? În ce fel constrânge ea posibilitățile de a ordona cuvintele într-o propoziție (în limba engleză, în cazul experimentelor lui Wanner și Marantsos)?

Să pornim de la sintagma următoare:

*To play a sonata on a piano.*

Înțelesul acestei sintagme este "a cânta o sonată la pian". Sintagme precum "to play" (care are înțelesul de "a cânta la un instrument" aici) au un fel de locuri goale (de "lacune" le vom zice pe scurt) care le însoțesc și care trebuie umplute pentru a înțelege despre ce este vorba. De exemplu, "to play \_" are o lacună marcată prin "\_" și, de asemenea, "on \_" are o lacună care dacă nu este umplută înțelesul rămâne suspendat (noi stăm parcă în așteptarea elementului care să ne arate la ce instrument se cântă).

Fie acum propoziția următoare:

(3) *A sonata is easy to play \_ on a piano.*



Sintagma "a sonata" a fost deplasată din poziția în care ar umple lacuna de după "to play". Ea trebuie reținută în memoria HOLD până în momentul în care va fi folosită pentru a umple lacuna respectivă.

Ce se întâmplă oare dacă sunt mai multe elemente de reținut în memoria HOLD? În ce structură vor fi dispuse ele? Pentru a testa acest lucru se poate recurge la întrebări cu "what" (ce).

O întrebare cu "what" este ușor de construit substituind, de pildă, pe "what" sintagmei "a sonata" în propoziția (3):

(3') *What is easy to play \_ on a piano?*

Ceea ce se obține nu este mulțumitor pentru că nu există cel puțin două elemente de reținut în memoria HOLD. Singurul element de reținut este chiar "what". Se poate modifica însă întrebarea în felul următor:

(4) *What is a sonata easy to play \_ on \_?*

Înțelesul întrebării (4) este, cu aproximație, "la ce instrument este o sonată ușor de cântat?". Se observă efortul cerut în limba engleză de cele două lacune. Ele trebuie umplute, în ordine, cu "a sonata" și "what"! Dacă HOLD are o structură de tip stivă, lucrul acesta decurge neproblematic: mai întâi este reținut "what", iar "a sonata" va sta, ca să spunem așa, undeva deasupra lui "what"; când elementele acestea sunt extrase din stivă, primul peste care nimerim este "a sonata" și abia apoi "what".

De ce am accepta însă ipoteza de mai sus? Poate structura memoriei HOLD este alta? Ceea ce ne face să acceptăm ideea stivei este ininteligibilitatea întrebării următoare:

(5\*) *What is a piano easy to play \_ on \_?*

Întrebarea (5\*) s-ar dori poate să fie doar o variantă a întrebării (3'). Deplasarea suferită în (5\*) de sintagma "a piano" încurcă însă lucrurile. Explicația, conform ipotezei lui Wanner și Marantsoos, este că se recurge la memoria HOLD și, în momentul extragerii, primul element peste care dăm este "a piano", pe

care îl folosim (fără succes!) pentru a umple lacuna de după "to play". Or, aceasta este structura de tip stivă.

Cei care se ocupă cu filosofia științei știu prea bine că ipotezele științifice sunt failibile. Ele nu sunt stabilite cu certitudine.

Nici ipoteza lui Wanner și Marantsos nu este desigur o excepție. Ea nu este oferită mai sus cu titlul de "adevăr despre memorie", ci ca o ilustrare a felului în care putem înainta în studiul memoriei. Din punct de vedere filosofic, aceste precizări sunt foarte semnificative.

Prin urmare, mai devreme sau mai târziu, trebuie să ne așteptăm ca o ipoteză să fie confruntată cu fapte care o contrazic. Așa, de pildă, ipoteza lui Wanner și Marantsos nu poate da direct socoteală de exemplul următor:

*John has his mother to consider.*

Propoziția aceasta inocentă spune ceva de genul că "John la mama sa trebuie să se gândească". Toată problema este că "to consider" are două lacune care trebuie umplute ( \_ to consider \_): când o umplem pe prima ne dăm seama cine anume are grijă; când o umplem pe a doua ne dăm seama de cine se are grijă. Atâta doar că "John", o dată pus în stivă, nu poate fi regăsit la momentul potrivit pentru a umple prima lacună (deoarece primul element în stivă este *his mother* "la mama sa"). Și totuși propoziția este inteligibilă!

Asemenea exemple dau întotdeauna bătaie de cap celor care făuresc și testează ipoteze. Ce fac cercetătorii în asemenea momente? O strategie ar fi abandonarea ipotezei. Altă strategie (mai frecventă !) este de a trata exemplele de genul celui de mai sus ca pe niște anomalii care pot fi explicate prin considerații speciale.

De pildă, se poate susține că prima lacună a lui "to consider" se efasează, umplerea ei nemaifiind luată în considerare. Fără alte analize, "John" este și subiectul lui "to consider".

Chiar dacă o procedură de genul celei de mai sus poate să ni se pară îndoielnică, ea este uzitată. Iar atâta timp cât nu se propune o ipoteză mai bună, ipoteza dată nu este abandonată.

Morala este că e oricum preferabil să lucrăm cu ipoteze. Acestea sunt miezul cercetării științifice, iar alternativa sunt certitudinile sprijinite pe o autoritate sau alta. De bine de rău, între ipoteze știm cum să alegem. Între autorități nu știm cum să alegem și singura soluție cu cap este toleranța: nimeni nu are autoritatea supremă, fiecare susține teza sa. Dar atunci ne trezim cu mai multe "certitudini"?! Desigur că nu; ar însemna să ne contrazicem. Așadar, această mică demonstrație ne arată că suntem condamnați să recurgem la ipoteze.

### Limbajul în perspectivă computațională

Limbajul este exemplul tipic de capacitate a minții studiată de către știința cogniției. Este, de altfel, liantul care unește parcă mai bine decât orice altceva disciplinele care compun știința cogniției. Capitolul de *lingvistică computațională* este de acum prezent în manuale de lingvistică generală.<sup>159</sup> Evident, prelucrarea limbajului natural este discutată în manualele de inteligență artificială.<sup>160</sup> Normal, limbajul și problemele acestuia preocupă de zeci de ani neuroștiința (localizare pe creier, afazie etc.) și psihologia (însușirea limbajului și alte probleme). În sfârșit, în ultimul secol cel puțin, interesul pentru limbaj al filosofiei a fost poate mai viu ca niciodată.

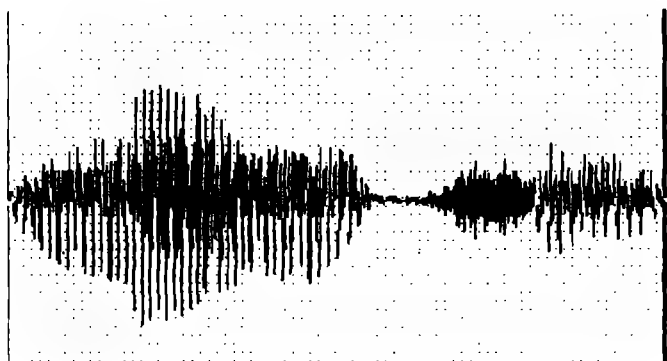
---

<sup>159</sup> Vezi, de exemplu, Judith Klavans, "Computational Linguistics" in William O'Grady, Michael Dobrovolsky (editori) și Mark Aronoff (editorul versiunii americane), *Contemporary Linguistics. An Introduction*, New York. St. Martin's Press, 1993, pp. 532-567.

<sup>160</sup> Vezi, de pildă, "Natural Language Processing", in Dean. Allen și Aloimonos, *Artificial Intelligence*, Menlo Park. Addison-Wesley, 1995, pp 489-537.

Studiul limbajului oferă, de asemenea, posibilitatea de a trece în revistă tot evantaiul de metode la care recurge abordarea computațională.

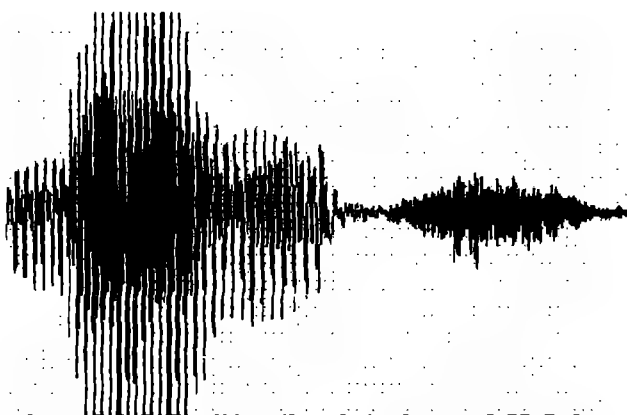
Ca și în cazul vederii, trebuie să începem cu ceea ce am denumit metaforic *rasterizare*. Înțelegerea vocii umane și posibilitatea de a răspunde vocal sunt acum la îndemână. Dacă aveți un computer conectat la Internet, puteți obține (gratis) o licență de folosire a unui sistem de dezvoltare de programe care analizează și sintetizează vorbirea vizitând pagina de web de la *Center for Spoken Language Understanding (CSLU)* de la *Oregon Graduate Institute for Science and Technology (OGI)*.<sup>161</sup> Veți găsi acolo cele mai variate unelte, inclusiv programele necesare pentru a realiza spectrograme (analize grafice ale semnalelor vocale). Undele sonore produse de vocea mea atunci când am pronunțat pe *minții* din sintagma "filosofia minții" arată ca în prima imagine. Evident, programul de computer este în stare să construiască nu doar imaginea, ci și să transforme totul în liste de biți care pot fi prelucrate ulterior și apoi iarăși transformate în sunete pe care le percepe urechea umană.



<sup>161</sup> La adresa <http://cslu.cse.ogi.edu/toolkit/>. De acolo puteți avea acces la documentație și, dacă aveți un computer foarte performant, puteți instala vastul sistem integrat (cu mii de programe) ce permite crearea de software care folosește limba naturală.

Dacă am spus simplu *minți*, spectrograma a arătat după cum se vede în cea de a doua imagine.

Deși exemplul pare banal, sensul său filosofic este mai profund decât am crede la prima vedere. Ce ilustrează el în fond? Dincolo de metafora *rasterului*, exemplul ilustrează mecanismul prin care un semnal oarecare din exterior capătă o *reprezentare* în computer, a cărei *formă* poate fi prelucrată. Noi nu prelucrăm unde sonore ca atare, așa cum aş putea-o face dacă aş modifica, să zicem, acustica camerei mele sau pe aceea a sălii de curs. Și totuși ceea ce rezultă, după ce este iarăși transformat




în undă sonoră, este perceptibil de către noi ca orice alt sunet.

*Mutatis mutandis*, aceleași considerente se pot formula și atunci când se recurge la recunoașterea optică a caracterelor. Efectiv, computerul poate "citi" textul. Lucrul acesta se poate face și cu un program de computer. De pildă, am tipărit pe printer textul unui program PROLOG și l-am prelucrat la facultate cu ajutorul scannerului. Fără a avea programul pe dischetă îl poți rula în acest fel. Totul începe și în cazul acestei "lecturi" cu un raster (de această dată chiar la propriu!) și se termină cu acțiuni efective ale computerului (de pildă, un dialog cu utilizatorul, precum în clasicul program "Eliza" al lui Weizenbaum).

În figura alăturată se vede un citat din *Cunoașterea limbii* de Noam Chomsky, pe care programul încearcă să-l

“recunoască” (mai precis, să transforme punctele de pe hârtie în coduri care pot fi folosite de către dispozitivele computerului). Se văd foarte bine în figură punctele rătăcite ale rasterului.

	2) (1)	P-*GN GV
(i)	P →  GN GV	
(ii)	GV → V GN	
(iii)	GN → DET N	
(iv)	GN → N	
(v)	V → <i>hit/lovi</i>	
(vi)	N → <i>boy/băiat</i>	
(vii)	N → <i>John/Ion</i>	
(viii)	DET → <i>the/arti</i>	
(ix)	X → ...	

Se observă cum programul nu poate găsi codul pentru săgeată. Greșeala este în partea de sus a figurii. În partea de jos este încercuit vârful de săgeată care a provocat neînțelegerea. Baza săgeții este luată drept o linie.

Să presupunem că totul merge însă, în această fază, fără probleme. Ceea ce urmează este prelucrarea codurilor obținute sau generarea unor coduri, utilizabile apoi pentru a produce pe linia de ieșire plus periferice de semnale (verbale sau scrise) care sunt percepute de către oameni ca mesaje într-un limbaj natural. Așadar, din ce se compune limbajul natural din punctul de

vedere al acestei etape a procesului computațional? Abstract vorbind, pe de o parte sunt categoriile în care sunt grupate "cuvintele", iar pe de altă parte sunt regulile (fie ca în imaginea de mai sus, ca la Chomsky, fie ca reguli de genul celor discutate aici în capitolele anterioare). Categoriile cuvintelor sunt de două feluri:<sup>162</sup> cuvinte cu conținut și cuvinte funcționale.

În fond, din punctul de vedere al procesului computațional, ceea ce contează este faptul că fiecărui element de limbaj (cuvânt, dacă vrei să-i spunei așa) îi este atașată o etichetă. Regulile arată apoi cum pot fi grupate cuvintele pentru a forma *sintagme* (sau, altfel spus, *grupuri* de cuvinte). La rândul lor sintagmele pot fi grupate și se dau reguli pentru a construi propoziții.

## Motorul semantic

John Haugeland scrie că termenul de *motor semantic* a fost propus de către Daniel Dennett.<sup>163</sup> Tot el explică în felul următor ideea de motor semantic: acesta este un sistem formal automat și astfel interpretat încât să aibă tot timpul sens. Este, cu alte cuvinte, o generalizare a ideii logice uzuale că mașinăria inferențelor trebuie să conserve adevărul, dar să permită efectuarea raționamentelor ținând cont doar de *forma* enunțurilor.

Termenul de "motor" folosit mai sus este o metaforă fericită. Un *motor*, în sens tehnic, "transformă o formă oarecare de energie în energie mecanică a unor corpuri solide în mișcare"<sup>164</sup>. Un "*motor semantic*" este alimentat cu secvențe care au înțeles și produce tot secvențe cu înțeles.

Dacă parcurgem cartea lui Chomsky *Cunoașterea limbii* vom da peste sistemul de reguli prezentat într-o figură mai

<sup>162</sup> Cf. Dean, Allen și Aloimonos, *Op. cit.*, p.491.

<sup>163</sup> John Haugeland, "Semantic Engines". in Haugeland (ed.), *Mind Design*.

p.24.

<sup>164</sup> *Dictionar politehnic*, București, Editura Tehnică, 1957, p.431.

sus.<sup>165</sup> Din punctul de vedere al mașinăriei formale a motorului semantic categoriile din care fac parte cuvintele cărora li se aplică regulile nu au semnificație.

Primul lucru pe care-l putem face cu ajutorul regulilor este să verificăm dacă un grup de cuvinte este sau nu corect format. Dacă ne este familiară logica, sensul acestor verificări ni se va părea evident. Nu întâmplător, cred, PROLOG pare limbajul cel mai nimerit pentru a ilustra această idee.

Să zicem că în clasa *numelor* (nume comune și nume proprii, în gramaticile tradiționale) este un singur membru, "Ion", iar în clasa *verbelor* sunt cuvinte precum "scrie" sau "citește". Atunci regulile gen Chomsky capătă în KPROLOG forma pe care o vedeți în figură.

```

Ce stiu?> (CONSECINTA (nume ion))
Ce stiu?> (CONSECINTA (verb scrie))
Ce stiu?> (CONSECINTA (verb citeste))
Ce stiu?> (CONSECINTA (grup_nominal X) DIN
> (nume X))
Ce stiu?> (CONSECINTA (grup_verbal Y) DIN
> (verb Y))

```

Acestea sunt reguli minime, ca să spunem așa. Din ele se vede că un grup nominal poate fi format și dintr-un singur nume. Cazul grupului verbal este similar. În română, ca și în multe alte limbi, o propoziție poate fi asamblată dintr-un grup nominal și un grup verbal, ceea ce și stipulează regula următoare.

```

Ce stiu?> (CONSECINTA (propozitie N V) DIN
> (grup_nominal N) (grup_verbal V))
Ce stiu?>

```

Până acum i-am spus noi computerului ce are *de știut*. Acum îl vom întreba ce știe.

<sup>165</sup> Vezi Noam Chomsky, *Cunoașterea limbii*, București, Editura Științifică, 1996, p.67.



```

Ce stii?> (CONSECINTA?
> (propozitie ion scrie))

```

Satisfacuta

```

Ce stii?> (CONSECINTA?
> (propozitie ion citeste))

```

Satisfacuta

Din câte știe sistemul, "scrie Ion" nu este o propoziție, după cum se vede din răspunsul la întrebarea pe care o punem în acest sens.

```

Ce stii?> (CONSECINTA?
> (propozitie scrie ion))

```

Mesatisfacuta

Se vede imediat, din exemplele de mai sus, că sistemul KPROLOG se folosește de "nume", "verb", "propoziție" ca de niște simboluri ca orice alte simboluri. Jocul este pur formal. Am fi putut pune orice alte simboluri acolo, atâta timp cât păstrăm constantă funcția (rolul) pe care-l joacă în sistem. Evident, nici cuvinte-cheie KPROLOG, precum "CONSECINTA", sunt folosite în același mod. Atâta doar că aici intră în joc considerente legate de scrierea interpretorului. Alegerea cuvintelor folosite nu are însă o importanță internă. În PROLOG-ul standard, construcții precum "CONSECINTA ... DIN ..." sunt, de altfel, simbolizate diferit.

Farmecul stilului de tip Lisp folosit de către Kamin este acela de a permite punerea în mod limpede în evidență a modului în care propozițiile sunt construite din structuri cuibărite (vezi figura în care se extinde ideea de grup verbal).

```

Ce stii?> (CONSECINTA (nume prostii))
Ce stii?> (CONSECINTA (grup_verbal (A B))
> DIN (verb A) (nume B))
Ce stii?> (CONSECINTA? (propozitie
> (ion (scrie prostii))))

```

Satisfacuta

Reprezentări de acest gen (cu paranteze cuiărite și etichete puse imediat după deschiderea parantezei) sunt utilizate și de către Chomsky.<sup>166</sup>

Gramaticile de genul celei de mai sus sunt programe. Cel de mai sus este scris în PROLOG și maschează poate unele caracteristici ale programelor scrise în limbaje procedurale. Oricum, categoriile gramaticale sunt structuri de date. Operațiile de bază sunt cele de asamblare a unor liste de cuvinte. Operațiile acestea, evident, pot fi repetate. Se pot stipula condițiile în care se face asamblarea și așa mai departe.

Programele de acest gen pot fi, de asemenea, folosite pentru a analiza gramatical texte. De această dată, programul va produce o structură care reflectă gramatica textului respectiv. Iar, până la urmă, va fi identificată și categoria din care face parte fiecare cuvânt.

Al treilea mod de utilizare a acestor programe este cel prin care sunt generate texte.

O sursă excelentă de sisteme sofisticate de a scrie astfel de programe o constituie cartea lui Terry Winograd *Language as a Cognitive Process*<sup>167</sup>. Winograd descrie, mai întâi, gramaticile bazate pe șabloane. Apoi sunt prezentate gramaticile transformaționale și gramaticile bazate pe rețele de tranziții. Acestea din urmă constituie miezul cărții. Winograd trece în revistă însă și alte gramatici, completând astfel tabloul variat al modurilor de analiză a limbilor naturale.

Probabil că multe cititoare sau cititori au auzit de numele lui Noam Chomsky. Când este vorba de inteligența artificială și analiza computațională a limbii naturale personalitatea al cărei nume ar trebui amintit cu prioritate este însă Terry Winograd.

<sup>166</sup> Chomsky, *Op.cit.*, passim.

<sup>167</sup> Terry Winograd, *Language as a Cognitive Process*, vol. I: Syntax, Reading (Mass.), Addison-Wesley, 1983, 640 p. După cum rezultă din pagina de web a lui Terry Winograd (<http://hci.stanford.edu/~winograd/publications.html>) el nu a publicat niciodată volumul care trebuia să fie consacrat semanticii. Probabil că motivele țin de cotitura în orientarea sa intelectuală semnalată în text.

Merită cu prisosință să amintim câteva repere extrem de interesante ale biografiei lui Terry Winograd. În 1972, Terry Winograd a devenit faimos pentru programul său capabil să execute comenzi formulate în limba engleză. Limita programului ține de faptul că orice dialog trebuie să se mărginească la o *microlume* (de pildă, la o cameră în care sunt corpuri geometrice variat colorate care trebuie manipulate de către program).<sup>168</sup> *Language as a Cognitive Process* marchează încheierea perioadei în care Winograd aderă la proiectul clasic al inteligenței artificiale. Filosofic vorbind, Winograd începe să respingă ideea unei semantici realiste. Obiectele depind de minte. Iar mintea este înrădăcinată în lume. Winograd cunoaște o cotitură<sup>169</sup> în concepția sa. El crede că și modul de a concepe programe de calculator trebuie să reflecte ideea că a ne făuri o unealtă înseamnă a proiecta un mod de viață.<sup>170</sup>

<sup>168</sup> Terry Winograd, *Understanding Natural Language*, New York. Academic Press, 1972, 191 p. (apărută și în *Cognitive Psychology* 3, 1/1972, pp. 1-191). Eu personal am aflat de carte prin intermediul traducerii în rusă (Moscova, Mir, 1972, 296 p.). Iată un exemplu de dialog între om și computer din cartea lui Winograd:

Omul: *Grasp the pyramid.*

Computerul: *I don't understand which pyramid you mean.*

Omul: *Find a block, which is taller than the one you are holding and put it into the box.* (ed. rusă, p.22)

<sup>169</sup> Vezi "Limits of the approach", în *Language as a Cognitive Process*, pp. 28-29. Winograd recunoaște meritele criticii (heideggeriano-wittgensteiniene) formulată de către Dreyfus (p. 32). Iar dacă ținem cont de faptul că una dintre principalele ținte ale criticii în prefața la ediția a doua din *What Computers Can't Do* este chiar Winograd, schimbarea este remarcabilă. Winograd spune că a publicat *Language as a Cognitive Process* conștient că există o incongruență între acest volum și cotitura intervenită în concepțiile sale cu privire la limbaj. Speranța sa era însă că, pe termen lung, materialul cu caracter tehnic din carte își va păstra utilitatea, dar se va integra în mod diferit într-o înțelegere a limbajului ca fenomen uman (cf. p. IX).

<sup>170</sup> Cartea care reprezintă vederile mai noi ale lui Winograd este scrisă împreună cu Fernando Flores (Winograd și Flores, *Understanding Computers and Cognition: A New Foundation for Design*, Norwood, NJ: Ablex, 1986, 220 p. Ediția paperback este publicată de Addison-Wesley, 1987). Ea a fost tradusă în franceză, germană, italiană, japoneză și spaniolă.

Winograd compară explicit perspectiva sa cu aceea a motorului semantic, așa cum este formulată de Haugeland.<sup>171</sup> După Winograd, drumul clasic de la sintaxă la semantică și de acolo la pragmatică trebuie inversat. De plecat trebuie plecat de la pragmatică. Nu forma limbajului este importantă, ci faptul că oamenii acționează cu ajutorul limbajului (Winograd îi menționează, în acest context, pe Austin și Searle). Actele de vorbire participă la un dans conversațional. Stările prin care trece conversația nu sunt stările mentale ale unuia sau altuia dintre participanți. Rețelele de tranziții de la o stare la alta nu mai servesc acum la reprezentarea unei gramatici. Programele construite de Winograd se centrează, după cotitura intervenită în concepția sa, pe conversație și opțiunile pe care le oferă aceasta.<sup>172</sup> În programele construite după cotitură nu există tentativa de a imita înțelegerea limbajului. Ele sunt unelte cu caracter general, tot așa – spune Winograd – cum un editor de texte permite să structurezi textul (în paragrafe etc.) dar nu lasă deschise opțiunile legate de conținut.

Opera lui Winograd, în ultimul deceniu al secolului al celui de al doilea mileniu, se centrează pe realizarea de programe care oferă un mediu de coordonare a acțiunilor celor care cooperează.<sup>173</sup> Proiectarea acestui mediu trebuie să țină cont, după Winograd, de relațiile dintre indivizi, de scopuri comune, dar și de conflicte, de autoritate și de modurile în care programul ca atare modifică rețeaua de relații interpersonale.<sup>174</sup>

---

<sup>171</sup> Terry Winograd, "A Language/Action Perspective on the Design of Cooperative Work", in *Human-Computer Interaction* 3, 1/1987-88, pp. 3-30. Versiune pe web: <http://hci.stanford.edu/~winograd/papers/language-action.html>

<sup>172</sup> Vezi meniurile programului "The Coordinator" în Winograd, *Op. cit.*

<sup>173</sup> Vezi, în acest sens, lista sa de publicații citată mai sus.

<sup>174</sup> Convingerea lui Winograd este că în faza lor inițială computerele au fost percepute și folosite îndeosebi ca unelte de calculat. De aici accentul pe aspectul *computațional* al computerelor. La sfârșit de mileniu, o dată cu Internet-ul, accentul s-a mutat pe *comunicare*. Fundamentală este proiectarea de programe care sunt optime din perspectiva interacțiunii lor cu utilizatorul. (cf. Terry Winograd, "Introducere", *Bringing Design to Software*, Addison-

Întorcându-ne acum la motorul semantic, vedem cum acesta este pus într-o perspectivă care-i circumscrie mult rolul. Atunci când forma nu mai are decât rostul de a face posibilă precizarea unor distincții utile în conversație este firesc ca și motorul semantic, a cărui funcție este, în fond, cea de a modifica forme, să aibă un rol diminuat.

Ajunsă în acest punct, dezbaterea din știința cogniției nu pare însă deloc încheiată. Diferitele părți aflate în dispute pot reveni cu noi argumente. De pildă, de ce nu am putea spera să găsim totuși un mod adecvat de a reprezenta și contextul în așa fel încât să fie și el luat în calcul în motorul semantic? Sau dacă are atâta greutate conversația, cum se face totuși că suntem în stare să ne susținem punctul *nostru* de vedere în cadrul ei? Nu joacă totuși *independența* minții un rol decisiv? Cum arată o lume lipsită de minți independente? Nu cumva pare mai degrabă lipsită cu totul de minți! Iar dacă mințile independente au atâta greutate nu înseamnă că trebuie să revenim și să ne concentrăm asupra lor!

### Provocarea ascunsă

Propria mea experiență de învățare a limbilor străine, dar și intuițiile rezultate din conversațiile cu cei care reușesc să stăpânească limbi străine, îmi sugerează o idee care poate părea foarte curioasă. Cele sau cei care știu bine o limbă străină au o cunoaștere foarte aproximativă a regulilor gramaticale în sens convențional. Bănuiesc că prima reacție venită "din public" ar merge pe linia ipotezei că regulile au fost transformate într-un mecanism care funcționează inconștient. Am să revin la ideea caracterului inconștient al mecanismului respectiv, dar pe moment vreau să focalizez analiza într-o direcție diferită. Dar sunt oare aceste reguli, fie ele și automatizate, inconștiente, cele convenționale ?

Primul lucru pe care-l știu din experiență este că e recomandabil să parcurgi mai multe gramatici sau manuale. Unghiurile de vedere diferite te vor ajuta să înțelegi mai bine despre ce este vorba. Pe scurt, concluzia mea este că fiecare își făurește mai degrabă propriul set de reguli. Cei care învață puțin limbi sau deloc încearcă să-și însușească regulile altora.

Tehnica intelectuală tradițională era cea a fișelor și a creionului. Ea este perfect transferabilă pe computer. Dar a ține simple însemnări pe computer nu pare un prea mare progres. O direcție mult mai tentantă este aceea de a construi propriul program. Cu alte cuvinte, ar trebui să modelezi propriul tău expert în domeniu. A scrie de la zero un astfel de program ar fi extrem de complicat. Din fericire, există "cochilii" gata pregătite, cum ar fi CLIPS, care pot fi "umplute" cu reguli.

Exemplul simplu care urmează explorează felul în care poate fi construit un astfel de sistem de notițe. Notițele sunt formate dintr-un pachet de programe care-ți permit să ridici treptat schele de pe care să construiești structura unei cunoașteri mai solide a limbii. Cazul luat în considerare este cel al limbii finlandeze.<sup>175</sup>

Din pricina unor probleme în funcționarea CLIPS, literele finlandeze *ä* și *ö* au fost înlocuite cu perechi convenționale *a'* *o'*. Pronunția lor este aproximativ cea a lui *a* din englezescul *cat* și a literei *o* cu umlaut din germană. Litera *y* se pronunță în finlandeză ca și *u* cu umlaut în germană (sau *u* din franceză).

Filosofia acestui paragraf este că fiecare s-ar putea simți îndemnat să noteze lucruri diferite. Mă voi mărgini, de aceea, la lucruri pe care, într-o formă sau alta, le facem cu toții. De pildă, încercăm să desprindem tipare recurente în limbă. Limitele acestei metode sunt foarte mari. Dar ea a funcționat cu succes în programe precum Eliza.

<sup>175</sup> Marea majoritate a informațiilor lingvistice provin din A. G. Morozova, Учебник финского языка, Petrozavodsk, Karelia, 1972, 288 p.

```
CLIPS> (load ".\\doi.clp")
```

```
*****
```

```
CLIPS> (reset)
```

```
CLIPS> (run)
```

```
Care este firul?
```

```
casa
```

```
Presupoziție ?
```

```
vecin obiect
```

```
ta'ma' on tala
```

```
5 rules fired
```

Ideea în fragmentul de mai sus este că tragem de un “fir” pentru a descoperi un șablon. Înțelesul șablonului găsit este “aceasta este o casă”. Programul nu furnizează însă asemenea traduceri. Prefer să trag de fire pentru a vedea alte exemple (pentru *on*, de pildă). Cuvintele românești sunt doar “capetele” firelor. *Presupozițiile* filtrează șabloanele. Este vorba deci de ceva aflat în vecinătate și care este etichetat drept obiect. Caseta următoare redă fragmente din textul programului scris în CLIPS.

```
(defrule casa
```

```
  (urma casa)
```

```
  =>
```

```
  (assert (ipoteza talo))
```

```
  (assert (vezi talo))
```

```
  (assert (presup-ip obiect)))
```

```
(defrule talo
```

```
  (vezi talo)
```

```
  (presup $?sup)
```

```
  (presup-ip $?hyp-sup)
```

```
  =>
```

```
  (if (and (eq (subset (mv-append obiect) $?hyp-sup) 1) (eq (subset (mv-append vecin obiect) $?sup) 1))
```

```
    then
```

```
      (assert (sablon ta'ma' on talo))
```

```
      (assert (cheie talo))
```

```
    else .....
```

```
(defrule carte
```

```
  (urma carte)
```

```
  =>
```

```
  (assert (ipoteza kirja))
```

```
  (assert (vezi kirja))
```

```
  (assert (presup-ip obiect)))
```

```
(defrule kirja
  (vezi kirja)
  =>
  (assert (vezi talo)))
```

Regulile permit mersul pe urma unui “fir”. Metafora este aceea a muncii de detectiv. Trebuie să folosești indicii (sau fire) pentru a rezolva problema.

Fiecare poate plasa într-o regulă ceea ce crede că ar caracteriza mai bine nodul respectiv al firului. Anumite puncte pot avea un caracter paradigmatic. Către ele trimit la alte fire. Ca *obiect*, o carte poate fi precum o casă: aproape, departe etc. Este posibil să introducem diferențe dacă este vorba despre carte ca sursă de informații ș.a.m.d.

Următorul mecanism este cel al înlocuirilor în șabloane, în limitele permise de presupozitii. Dacă știu că se poate spune *seina' on valkoinen* (peretele este alb), atunci pot schimba culoarea în *musta* (negru).

Finlandeza are un sistem bogat de cazuri (restrângs în programele de mai sus la douăsprezece). Iată cum putem merge

```
Fir?
cazuri:
Presupozitii?
static intern stare
Tiparul sugerat este: ja'rv'i on ja'a'ssa'
```

pe urmele cazurilor.

Șablonul propus s-ar traduce în românește prin “lacul este înghețat”. Atâta doar că în finlandeză se spune *lacul este ...* (cu “gheață” pus la cazul zis *inesiv*). Ce se sugerează mai sus este că pot să construiesc propriul meu mod de a merge pe firul cazurilor sau al altor forme ale cuvintelor finlandeze.

Dacă am fi utilizat lista de presupozitii “intern static activitate”, am fi găsit un exemplu de genul *ha'n on kalassa*, care înseamnă “el/ea este la pescuit” (*kala* înseamnă “pește”, iar *kalassa* este *inesivul*). Este ca și cum am spune ceva de genul



“el/ea este la pește”. Tot așa, se spune “el/ea este la muncă” (muncește) folosind inesivul de la *tyo*’.

După cum se vede, mai ales sistemul unei limbi suficient de depărtate de cea de acasă din punct de vedere lingvistic permite variate moduri de elaborare a regulilor. Fiecare prinde diferit “firele”. Gândiți-vă, de pildă, că finlandeza, pentru a exprima pe “nu este”, are un verb. Englezii ar putea merge pe firul lui “do not” (atenție însă că această construcție nu se folosește în cazul lui *is* “este”, a cărui negație este *is not*).

Alte trăsături ale limbii care pot fi prinse în mod diferit, în funcție de punctul de plecare, sunt lipsa unui verb “a avea” sau sufixele posesive atașate substantivelor (*kissamme on kaunis* înseamnă “pisica noastră este frumoasă”). Sunt, de asemenea, numeroase capcane la folosirea cazurilor pe lângă verbe.

Morala este deci că mașinăria poate fi personalizată. Idealul inteligenței artificiale clasice ar fi ca programul să fie capabil să preia orice date (cuvinte, distincții între categorii, presupoziii) în limbaj natural și să creeze automat un program de genul celui de mai sus, care să-mi permită să regăsesc informații și să folosesc reguli pentru a înțelege și pentru a construi un text.

Care ar fi însă provocarea ascunsă în discuția de mai sus? Computerele au fost la început percepute ca fantastice mijloace de calcul; apoi a venit rândul prelucrării masive de date nenumărate, bibliotecile s-au mutat pe CD-uri, arhivele au devenit electronice; apoi a venit momentul comunicării masive pe Internet, rețelele de computere tind să ia locul telefoanelor, radiourilor și televizoarelor, comerțul, bursele chiar se mută pe web. S-ar părea însă că, din punctul de vedere al minții umane, nu suntem însă departe de punctul în care am fost înainte. Diferența dintre a așterne gândurile pe hârtie și a le pune în fereastra unui editor nu este dramatică. Poți face mai ușor trimiteri, poți tehnoredacta textul. Dar acestea erau oricum operații tradiționale. *A scrie* însă programe este o nouă provocare. Șansa este aceea de a lăsa în urma ta în lume nu doar propoziții, ci și moduri de a structura date, opțiuni între diversele căi de

prelucrare a datelor și de producere de noi distincții și așa mai departe.

Provocarea aceasta ne dă o idee destul de precisă cu privire la ce am putea face. Așa cum arătata Winograd în cazul comunicării, programul inițial este o cochilie care urmează să fie completată.

Dacă este vorba de descoperirea limbilor străine, cochiliile pe care le știu eu sunt totuși prea dependente de limba de acasă. Iar principalul serviciu este acela de a te ajuta să creezi liste de cuvinte. Problema este că o limbă suficient de diferită corespunde prost cu cea de acasă. Distincțiile gramaticale sunt altele și lingviștii au constatat de multă vreme că ele trebuie reconstituite din interiorul limbii. De aici imposibilitatea de a avea un program gata pregătit foarte elaborat. Esențialul trebuie și este bine să fie făcut cu propria minte. Rezultatul este ar putea fi acela de a avea *amprenta unor minți*.

### **În loc de rezumat: și "mașină" (de calcul) este tot o metaforă**

Critici și apărători ai abordării cognitive a minții împărtășesc o eroare subtilă. Computerele (nu doar cele reale, ci și construcțiile imaginare) – spun criticii – nu sunt decât *mașini*. Ele sunt unelte care ar putea fi utile în studiul minții, dar nu mai mult de atât.

Pe de altă parte, apărătorii computaționaliști ai cognitivismului sunt fericiți să constate că un computer poate fi *un sistem formal*. Alimentat cu reprezentări adecvate, un asemenea sistem ar reproduce funcționarea unei minți.

Computerele, așa cum le-a imaginat Alan Turing, pot fi desigur și sisteme formale. Dar pot fi și orice alt sistem.

În cazul computerelor reale intră firește în joc limitele materialelor din care sunt făcute și ale programelor pe care le rulează. Ceea ce există potențial este acea capacitate pe care nu o are ceasul cu arc și roți dințate: posibilitatea de realizare a unei schimbări în funcționare care să-l transforme în orice

altceva. De aici caracterul metaforic al termenului de "mașină". Aceasta este comparația aflată cel mai la îndemână.

Pe scurt, ceea ce a caracterizat abordările computaționale ale minții prezentate până acum este convingerea că rezultatele valoroase vor veni din transformarea "mașinii" într-un sistem formal, într-o mașinărie perfectă de manevrat simboluri ținând cont doar de forma lor. Demersul acesta are valoarea sa, dar și limite ce țin de lipsa unei înrădăcinări adecvate în lume.

Este oare posibil ca una dintre limite să fie și existența unor *sisteme inimitabile*? Ar fi aceste sisteme mințile umane? Aceasta va fi tema penultimului capitol, dar înainte de a ajunge la ea vom continua să explorăm beneficiile pe care ni le aduce abordarea computațională.

## Modularitatea minții

În capitolul al treilea am arătat de ce un program organizat modular este superior unuia nemodularizat. Speculativ, ne-am putea întreba dacă o astfel de superioritate s-ar traduce și printr-un avantaj în cadrul procesului de evoluție a speciilor.

Jerry Fodor este cel care a susținut că există *dovezi* în sprijinul ideii că mintea umană are un caracter modular.<sup>176</sup> Cu alte cuvinte, psihologia facultăților mentale ar trebui să beneficieze de mult mai mult credit decât s-a bucurat în cea mai mare parte a secolului al XX-lea.

Fodor distinge patru moduri de a concepe *structurile mentale*. Primul mod este cel neocartezian, promovat de către școala lui Noam Chomsky.<sup>177</sup> Pentru neocartezieni structurile minții sunt bogate și diverse. Metafora favorită a lui Chomsky este cea a "organului". Așa cum avem mâini, avem și *organe mentale*. Ca și mâinile, aceste organe se dezvoltă pe baza unui

---

<sup>176</sup> Jerry A. Fodor, *The Modularity of Mind. An Essay on Faculty Psychology*, MIT Press, 1983.

<sup>177</sup> Fodor, *Modularity*, p. 3.

program genetic. Spre deosebire însă de mâini, ceea ce este înăscut în cazul organelor mentale este cunoașterea de care dispunem (pentru a putea, de pildă, vorbi o limbă).

Din cea de a doua perspectivă identificată de către Fodor, structurile mentale se constituie sub forma unei arhitecturi de facultăți orizontale.<sup>178</sup> Aceste facultăți sunt identificate *funcțional*. Criteriul îl constituie efectele pe care le produc. Sistemul memoriei, de exemplu, este identificat prin funcțiile sale. Tot așa și facultatea de judecare etc. Acestor facultăți le este accesibil orice conținut mental. Ele deserves, cu alte cuvinte, multiple domenii.

Cea de a treia perspectivă este cea a facultăților verticale.<sup>179</sup> Din acest punct de vedere nu ar exista atenție sau voință ori o facultate de judecare. Există mănunchiuri de competențe sau înclinații. *Talentul* pentru ceva (muzică, să spunem) ar fi un exemplu în acest sens.

Cea de a patra perspectivă identificată de către Fodor este cea asociaționistă. Pentru asociaționiști relațiile dintre conținuturile mentale sunt mecanice.<sup>180</sup> Dinamica mecanică a acestor conținuturi (atrakția sau respingerea ideilor) joacă un rol cheie.

Fodor precizează că există un contrast între asociaționism și cognitivism. Pentru cognitiști ceea ce contează sunt computațiile care pot fi făcute cu conținuturile mentale. Structurile mentale, din perspectivă computațională, pot avea un grad arbitrar de complexitate.<sup>181</sup>

Ceea ce-l interesează, până la urmă, pe Fodor sunt *sistemele cognitive modulare*. Sistemele cognitive sunt capabile să înregistreze ceea ce se întâmplă în afara lor. Orice mecanism capabil să *covarieze* cu evenimentele din mediu va furniza

---

<sup>178</sup> Fodor, *Modularity*, p. 10.

<sup>179</sup> Fodor, *Modularity*, p. 14.

<sup>180</sup> Fodor, *Modularity*, p. 31.

<sup>181</sup> Fodor, *Modularity*, p. 34.

sistemului *reprezentări* ale stărilor mediului.<sup>182</sup> Aceste reprezentări sunt apoi prelucrate computațional.

Sistemele cognitive sunt divizate, la rândul lor, în subsisteme. Fodor distinge clase mari de asemenea subsisteme. Recurgând la metafora folosită aici, prima clasă este formată din subsisteme care produc "raster". O altă clasă este cea a subsistemelor care constituie intrările în sistemul cognitiv. De asemenea, Fodor distinge și clasa procesorilor centrali.<sup>183</sup>

Este foarte interesant să ne oprim puțin asupra sistemelor care asigură intrările. Sistemele de intrare sunt sistemele percepției plus limbajul.<sup>184</sup> Fodor ilustrează astfel foarte bine o *teorie comunicațională a limbajului natural*.

Acum nu rămâne decât să constatăm că sistemele de intrare sunt module. La rândul lor, aceste sisteme, cum am aflat în cazul teoriei vederii, sunt modularizate.

Ca și în cazul programelor, o dată declanșate, modulele nu pot fi influențate în funcționarea lor. Dacă soția a deschis televizorul, oricât m-aș concentra asupra textului lui Fodor, tot aud știrile de la jurnal.<sup>185</sup> Se observă, de altfel, că intră în funcțiune și mecanismul de înțelegere a cuvintelor. Desigur, pot la un alt nivel să nu procesez fluxul de date care-mi parvine, dar de venit el tot va veni.

Modulele sunt *încapsulate informațional*: chiar dacă dau un rezultat greșit acesta tot nu poate fi corectat. Testul în acest sens îl oferă iluziile vizuale.<sup>186</sup>

Modularitatea în sensul lui Fodor are un caracter funcțional. De aceea, nu este obligatoriu să existe o localizare pe

---

<sup>182</sup> Fodor, *Modularity*, p. 39.

<sup>183</sup> Fodor, *Modularity*, p. 41.

<sup>184</sup> Fodor, *Modularity*, p. 44.

<sup>185</sup> Vezi Fodor, *Modularity*, pp. 52-53.

<sup>186</sup> Fodor, *Modularity*, p. 66.

creier a modulelor.<sup>187</sup> Procesarea centrală integrează datele furnizate de către sistemele de intrare.<sup>188</sup>

Una dintre concluziile cele mai provocatoare trase de către Fodor din modularitate este legată de ceea ce el numește *impenetrabilitate cognitivă*. Modulele nu pot fi influențate sau condiționate din exterior. Ideea este plauzibilă dacă ne gândim, din nou, la programe. Un program ar deveni extrem de încălzit și treptat s-ar prăbuși din pricina erorilor dacă nu ar fi modularizat. Modulele nu preiau tot timpul noi date sau, mai rău, noi instrucțiuni. Ele sunt doar apelate. Dacă și mintea este organizată modular, atunci sisteme precum vederea sunt cognitiv impenetrabile. Aceasta contrazice frontal ideea, populară în filosofia actuală, a influenței contextului, crezămintelor sau teoriilor asupra percepțiilor.

Credința în “încărcarea” de un fel sau altul a vederii sau altor sisteme de intrare este răspândită în filosofia științei și științele sociale. Acest contrast între cognitivism și modele standard în științele sociale nu este singular. Științele sociale sunt influențate de Whorf și modelul lor este cel al unei interpenetrări între cultură, gândire și limbaj. Cognitiviștii gen Fodor tratează limbajul ca pe un modul aparte.

### **Cum de au gândurile efecte în lumea fizică?**

Pentru mentaliști a fost foarte greu să explice eficacitatea cauzală a minții în lumea fizică. Pe de o parte, această eficacitate este un fapt. Dacă mă hotărâsc să șterg acest paragraf din carte, atunci el dispare fizic. Nu va mai exista imaginea sa pe ecran. Nu va exista nici pe foaia de hârtie. Pe de altă parte, lumea fizică și lumea mentală sunt distincte. A spune că “interacționează” între ele sau că una o condiționează pe cealaltă, fără a arăta care ar fi mecanismul acestor influențe, nu este decât o probă de

<sup>187</sup> Fodor, *Modularity*, p. 98.

<sup>188</sup> Fodor, *Modularity*, p. 103.

verbalism. În fond, știam deja că interacțiunea se produce. Ne interesează cum anume are ea loc.

Unul dintre motivele atracției abordării computaționale ține de capacitatea de a furniza o dezlegare problemei de mai sus. Această rezolvare este inteligibilă dacă ne reamintim ideea de *limbaj-mașină*.

Articularea teoretică a soluției o putem găsi în ipoteza lui Jerry Fodor privitoare la *limbajul gândirii*.<sup>189</sup> În ciuda “vechimei” sale și a faptului că Fodor însuși a schimbat recent argumentarea, ipoteza lui Fodor este în centrul multor dezbateri<sup>190</sup> și, mai important, se mulează pe firul argumentului pe care vreau să-l construiesc în această carte.

Punctul de start al demersului lui Fodor este teoria reprezentării. Computațiile au drept intrări reprezentări. De ce avem nevoie de reprezentări? De ce nu le-am respinge, după sfatul behaviorist, pe motiv că sunt entități mentale? Aici momente ale răspunsul lui Fodor sunt deosebit de pertinente: reprezentările ne permit să distingem *stări de lucruri* diferite.<sup>191</sup> În acest fel ne putem reprezenta și stări de lucruri *posibile*, nu doar pe cele reale. Importanța acestor stări posibile nu mai trebuie să fie, cred, subliniată.

Când este vorba despre reprezentări, strategia intelectuală este, în fond, aceeași cu cea pe care o vedem în programare atunci când se vorbește despre *structuri de date*.

---

<sup>189</sup> Jerry A. Fodor, *The Language of Thought*, Harvard University Press, 1975.

<sup>190</sup> Știința cogniției are o evoluție foarte rapidă. Rezultatele se pot învechi de la un an la altul. De aceea, teorii ca ale lui Marr despre vedere sau Fodor despre limbajul gândirii sunt teorii clasice. Teorii atât de vechi au mari șanse să fie depășite. Există însă și excepții. Criteriul de bază aici cred că este persistența unor dezbateri animate în jurul teoriei respective. La aceasta se adaugă, desigur, lipsa unor soluții alternative. Pentru literatura la zi referitoare la limbajul gândirii se poate consulta Larry Kaye, “The Language of Thought: Annotated Bibliography” din ghidul de filosofia minții al Universității Roma 3, consacrată ipotezei lui Fodor:

<http://www.uniroma3.it/kant/field/lotfbiblio.html> [accesat pe data de 5.11.1999].

<sup>191</sup> Fodor, *Language of Thought*, p. 33.

Computațiile sistemelor artificiale au drept intrări date de un anumit tip (numere, liste etc.). Tot așa mintea prelucrează date de un tip sau altul.

O dată ce admiți ideea reprezentărilor, ai nevoie de un limbaj în care să fie exprimate aceste reprezentări. Argumentul lui Fodor este că acest limbaj nu poate fi unul dintre limbajele zise naturale (limbajele folosite în procesul de comunicare). Cum am putea explica atunci gândirea copiilor mici sau a animalelor care nu cunosc nici o limbă naturală?<sup>192</sup> Trebuie să admitem, crede Fodor, că acesta este un *limbaj al gândirii*. Limbajul gândirii este unul singur pentru toți oamenii, în vreme ce limbile naturale sunt numeroase.

Fodor se concentrează, în cartea mai veche, asupra aspectelor de ordin logic ale limbajului gândirii. Asupra lor el avea însă să revină relativ recent cu noi idei (privitoare la concepte și felul în care ele sunt prezente în mintea noastră). Din punctul de vedere adoptat aici (care favorizează masiv tehnicile programării) ceea ce este semnificativ în limbajul gândirii este posibilitatea de a exprima în acest limbaj *reguli* (în sensul tehnic precizat anterior). Vom vedea în paragraful următor efectele lor în viziunea lui Fodor. Prezența regulilor asigură posibilitatea de a prelucra și reprelucra reprezentările. De aici și capacitatea minții de a obține un număr nelimitat de gânduri.

Legătura cea mai interesantă este însă cea cu limbajul-mașină.<sup>193</sup> Limbajul-mașină conține coduri de două feluri: date care urmează să fie preluate ca intrări de dispozitivele sistemului și coduri de activare a dispozitivelor. A plasa un cod de activare într-un dispozitiv este o operație de genul celei pe care o facem când apăsăm un întrerupător. Pozițiile deschis-închis pot, de altfel, fi interpretate ca "0" și "1". Dacă ne închipuim că un cod are opt biți, atunci opt întrerupătoare vor fi o reprezentare fizică potrivită. În funcție de felul în care sunt

<sup>192</sup> V. Fodor, *Language of Thought*, în special p.56 ș.u.

<sup>193</sup> Fodor prezintă lucrurile din perspectiva a ceea ce el numește proprietatea critică a limbajului-mașină de a avea formulele în corespondență de unu la unu cu stări fizice relevante computațional ale sistemului (cf. Fodor, *Language of Thought*, p. 67).



deschise sau închise, se înscrie un cod sau altul. Această operație de înscriere a codului este însă un eveniment fizic și declanșează un lanț de evenimente fizice. Concluzia este că o configurație sau alta a codurilor are efecte fizice diferite.

Astfel este deschisă calea pentru a produce o rezolvare a problemei cu care am deschis paragraful. Pe de o parte, limbajul gândirii are eficacitate cauzală după un tipar analog celui folosit în cazul limbajului-mașină. Semnale adecvate declanșează dispozitive. Pe de altă parte, putem explica și de ce mentalul nu face parte din lumca fizică. Lanțurile cauzale fizice pot evident modifica semnalele (codurile) limbajului gândirii, dar nu pot reconfigura sistematic semnalele limbajului gândirii. Schimbarea configurațiilor este operată de către sistem în baza examinării configurațiilor codurilor formulate în limbajul gândirii, nu a legilor fizicii.

Argumentul de mai sus este susținut evident de faptul că procesele computaționale nu sunt reductibile la legi fizice, dar au eficacitate cauzală. De aici, probabil, fascinația pe care o exercită.

La urmă, dar nu în cele din urmă, ar trebui menționat faptul că ideile lui Fodor intră (sau par să intre) în conflict cu critica lui Wittgenstein la adresa unui limbaj privat.<sup>194</sup> Crucial este însă, după Fodor, faptul că un limbaj folosit pentru computații (precum limbajul-mașină sau limbajul gândirii) nu presupune capacitatea cuiva de a *stabili* dacă termenii sunt folosiți consistent; el presupune doar ca ei *să fie* folosiți corect.<sup>195</sup>

Cred că Fodor are dreptate. Mie cel mai periculos mi s-ar părea argumentul lui Wittgenstein în interpretarea lui Kripke. De aceea l-am și prezentat în cartea de față. În versiunea lui Kripke punctul crucial ar fi stabilirea regulii care se aplică. Aici sunt însă două observații. Mai întâi, după cum arată Fodor, nici nu este nevoie să se stabilească nimic. A doua privește faptul că, în

<sup>194</sup> Fodor însuși discută ideile lui Wittgenstein (cf. Fodor, *Language of Thought*, în special argumentul de la p. 69).

<sup>195</sup> Fodor, *Language of Thought*, p.70.

cazul operațiilor de bază, codul acestora declanșează un dispozitiv. Astfel, întrebarea lui Kripke va suferi o deplasare și va privi funcționarea dispozitivului. Chiar dacă noi nu putem preciza ce anume va face dispozitivul (la urma urmei, funcționarea sa depinde de legi empirice!) efectele acestei funcționări, fie și indirect, le putem observa în acțiuni.

## Conexionism contra computaționalism

Dezbaterea anterioară este departe de a se fi încheiat. Eficacitatea cauzală a minții ar putea primi și alte explicații. Cea mai plauzibilă alternativă vine din direcția conexionismului. Un sistem conexionist, după cum am arătat, are o arhitectură bazată pe rețele de neuroni artificiali.

Într-un articol celebru, Paul Smolensky a oferit ceea ce el numește PTC (*Proper Treatment of Conexionism*) - o prezentare a conexionismului menită să arate superioritatea acestuia față de modelele zise clasice sau computaționaliste, de genul celui la care recurge Fodor pentru a formula ipoteza limbajului gândirii.<sup>196</sup> Smolensky pune modelele clasice sub semnul unei *paradigme simbolice*.<sup>197</sup> Termenul *simbol* este menit să accentueze două lucruri: computaționalismul folosește limbaje de tip Lisp<sup>198</sup>; simbolurile au un înțeles, în sensul că reprezintă obiecte externe sistemului.

---

<sup>196</sup> Paul Smolensky, "On the Proper Treatment of Connectionism". *Behavioral and Brain Sciences* 11, 1988, pp. 1-23. Reprodus parțial în Alvin I. Goldman, *Readings in Philosophy and Cognitive Science*, MIT Press, 1993, pp. 769-799.

<sup>197</sup> Smolensky, PTC, p. 773.

<sup>198</sup> Dacă examinați atent limbajele de tip Lisp prezentate în această carte veți vedea o oarecare imprecizie în această formulare: în Lisp structura de date centrală este *lista* (de simboluri). De asemenea, cred că structurile de date sunt mai bogate în aceste limbaje decât sugerează termenul ales de Smolensky. Ca și mulți alți autori, el pare însă atras de ideea că totul nu este decât o transpunere a unor idei familiare din logică. Adevărul este că logica e doar o parte a poveștii și, după cum am arătat, nu cred că este în centrul ei.

După Smolensky, conexionismul oferă modele care operează cu *subsimboluri*. Acestea sunt activitățile unui neuron dintr-o rețea conexionistă (vezi capitolul al treilea). Se poate opera cu *reprezentări* la acest nivel al subsimbolurilor.<sup>199</sup> Conceptul de *mașină virtuală* poate fi reintrodus; din perspectivă conexionistă este vorba despre mașini virtuale care se constituie peste acest nivel subsimbolic.<sup>200</sup>

Din perspectiva adoptată în această carte, cea a unor sisteme de dispozitive abstracte, sensul conexionismului cred că ar putea fi descifrat în felul următor: în computaționalism, pentru a realiza un dispozitiv virtual, se recurge la plasarea de coduri adecvate în dispozitive deja existente, cu alte cuvinte la un program în limbaj-mașină; în schimb, într-o rețea conexionistă se modifică ponderile conexiunilor și astfel apar, de fapt, noi dispozitive.

Se observă imediat, din prezentarea de mai sus, că miza disputei este legată de rolul limbajului. Conexionismul se poate plasa la un nivel la care nu trebuie recurs la un limbaj-mașină și nici la un program în sens computaționalist. Desigur că această abordare ar reteza temeiurile ipotezei limbajului gândirii.

Smolensky crede că PTC răspunde argumentelor cu care Fodor și Zenon Pylyshyn au încercat să pareze ascensiunea conexionismului. Nu este, de aceea, lipsit de interes să vedem ce spun cei doi computaționaliști.<sup>201</sup>

În esență, Fodor și Pylyshyn susțin că există o serie de *trăsături ale cogniției* care nu pot fi reproduse într-un sistem conexionist. Vom examina aici doar două dintre aceste trăsături.

Sistemele cognitive, spun Fodor și Pylyshyn, sunt *productive*. Termenul "productiv" este ușor de explicat dacă ne

---

<sup>199</sup> Smolensky, PTC, pp. 783-785.

<sup>200</sup> Smolensky, PTC, p. 790.

<sup>201</sup> Jerry A. Fodor și Zenon W. Pylyshyn, "Connectionism and cognitive architecture: A critical analysis", *Cognition*, 28, 1988, pp. 3-71. Articolul este reprodus parțial în Goldman, *Op.cit.*, pp. 801-818. Fodor și Pylyshyn au avut acces la articolul lui Smolensky, ca și Smolensky la articolul lor. Aceasta face, de altfel, cu atât mai interesante articolele respective.

gândim că în literatura de specialitate *regulilor* li se spune curent *reguli de producție* sau chiar *producții*. Ceea ce vor să spună Fodor și Pylyshyn este că un sistem care poate opera cu reguli (de producție) adecvate va genera, cu ajutorul unui număr finit de reguli, o infinitate de construcții (asamblate urmând regulile). În particular, gândirea umană este productivă.<sup>202</sup>

O altă trăsătură a sistemelor cognitive este *sistematicitatea*. Pentru a înțelege despre ce este vorba, Fodor și Pylyshyn ne propun un contrast între două moduri de a învăța o limbă străină: (1) memorarea expresiilor dintr-un ghid de conversație uriaș; (2) învățarea limbii în mod normal.<sup>203</sup> Dacă ne uităm la exemplul învățării limbii finlandeze, discutat de noi mai sus, vedem imediat în ce capcană formidabilă am intra dacă am învăța pe de rost șabloane corespunzătoare unor expresii românești. Paralelismul de structură dintre "el este la pescuit" și "lacul este înghețat" din expresiile finlandeze corespunzătoare ne-ar scăpa. Am avea în memorie doar simple liste de cuvinte.

Sistematicitatea este legată, până la urmă, tot de reguli, de regulile care ne permit să trecem într-o limbă de la o expresie la alta și, mai profund, de la cunoaștere explicită la cunoaștere implicită. Nu ajunge să memorăm liste pentru a ști o limbă, trebuie să stăpânim reguli.

Fodor și Pylyshyn reproșează, în fond, conexionismului că tratează reprezentările ca pe niște liste.<sup>204</sup> Doar cu liste nu poți face mare lucru. Este ca și cum în KPROLOG (v. mai sus) am avea doar construcțiile cu "(CONSECINTA (...))", dar nu și pe cele de forma "(CONSECINTA ... DIN ...)". Concluzia este că nimic, în afara intervențiilor *ad hoc* ale constructorului, nu ar putea determina formarea într-o arhitectură conexionistă a mașinilor virtuale necesare pentru a avea productivitate și sistematicitate. Trebuie să recurgem la un limbaj și mecanismele sale pentru a explica trăsăturile cogniției. De aici și ideea că

<sup>202</sup> Fodor și Pylyshyn, *Connectionism*, p. 33 ș.u.

<sup>203</sup> Fodor și Pylyshyn, *Connectionism*, p. 37.

<sup>204</sup> Fodor și Pylyshyn, *Connectionism*, p. 49.

argumentele lui Fodor și Pylyshyn se constituie într-un sprijin pentru ipoteza limbajului gândirii.

Care este replica lui Smolensky? Obiecțiile lui Fodor și Pylyshyn vizează doar anumite sisteme conexioniste, spune el. PTC include ipoteza unui *sistem dinamic conexionist* (v. capitolul al treilea pentru sisteme dinamice). Aceasta ar fi alternativa la mecanismele sintactice din paradigma simbolică. Smolensky crede că poate oferi, de asemenea, o alternativă și la semantica din paradigma simbolică.<sup>205</sup>

### Competiția ipotezelor

Știința cogniției nu este o "știință" doar cu numele. Ca în orice știință, există ipoteze alternative. Tot ca în orice știință, este hazardat să alegi între ipoteze pe baza unor argumente pur speculative.

Am mai putea adăuga la aceste observații că selecția ipotezelor este, de asemenea, greu de făcut pe baza unor rezultate de moment. Se confruntă, de fapt, *programe de cercetare*.

Nu are rost, din aceste motive, să tragem concluzii pripite cu privire la superioritatea unei ipoteze sau a alteia. Conexionismul a părut complet ieșit din competiție în anii 1960 din pricina incapacității de a modela calcularea unor funcții precum XOR. A revenit însă în anii 1980. Nu doar pentru că a modelat pe XOR. Conexioniștii au putut trece, între altele, peste critica lui Fodor și Pylyshyn pentru că au realizat rețele care modelează regulile de producție (vezi capitolul al treilea).

Orice rezultat practic obținut este însă sub semnul întrebării privitoare la limitele sale. Până la ce nivel de complexitate pot ajunge ierarhii de mașini virtuale din arhitectura rețelelor conexioniste? Eu personal bănuiesc că există unele bariere.

La urmă, dar nu în cele din urmă, trebuie trasă și concluzia că, în vreme ce în programele de cercetare din științele naturii filosofia este prezentă la nivelul *presupozițiilor*

---

<sup>205</sup> Smolensky, PTC, în special pp. 780-781.

programului, în știința cogniției filosofia este o parte a jocului propriu-zis. Și, pentru că am început capitolul cu un citat dintr-o lucrare a unui reprezentant marcant al computaționalismului, trebuie să menționăm în încheiere și interesul pentru legătura dintre conexiionism și filosofia minții.<sup>206</sup>

---

<sup>206</sup> Vezi, de pildă, Terence Horgan și John Tieson, *Connectionism and the Philosophy of Mind*, Kluwer Academic Publishers (în special William Bechtel, "Connectionism and the philosophy of mind: an overview", pp. 30-59).

## Noi orizonturi

Știința cogniției, după cum arătam în primul capitol, are un caracter interdisciplinar. Până acum interferențele explorate au fost însă doar cele dintre sistemele computaționale, filosofie și, în parte, lingvistică. Acest lucru este firesc dacă ținem cont de impactul masiv al inteligenței computaționale asupra științei cogniției în ansamblul ei. Pasul făcut în continuare este cel al examinării și altor interfețe, între filosofie și psihologie, între filosofie și teoria evoluționistă.

Interferențele de care vorbeam sunt *noile orizonturi* ale cooperării, inclusiv între discipline care păreau despărțite pe vecie, precum psihologia și filosofia.

## Psihologia cognitivă

Studiul capacităților cognitive ale minții umane (percepția, memoria, limbajul etc.) a existat și ar exista și independent de cognitivism. De altfel, cel care a sintetizat pentru prima oară perspectiva cognitivă în psihologia cognitivă. Ulric Neisser,<sup>207</sup> a renunțat ulterior la cognitivism, în favoarea unei abordări *ecologice*.<sup>208</sup> Conform vederilor sale mai târziu, centrală pentru percepție este activitatea subiectului. Informațiile dintr-o scenă pe care o percepe subiectul uman sunt infinite de bogate, iar omul recurge la un proces de selecție pozitivă.<sup>209</sup> Atenția nu este altceva decât percepție, din această

---

<sup>207</sup> Ulric Neisser, *Cognitive Psychology*, New York, Appleton, 1967. Cartea aceasta a contribuit, se pare considerabil, la răspândirea cognitivismului printre psihologi.

<sup>208</sup> A se vedea Ulric Neisser, *Cognition and Reality. Principles and Implications of Cognitive Psychology*, San Francisco, Freeman, 1976 (tradusă în rusă în 1981). Cartea este dedicată, de altfel, lui James și Eleanor Gibson, reprezentanți cunoscuți ai orientării ecologice.

<sup>209</sup> Neisser, *Cognition and Reality*, cap.5, p. 97 (ediția rusă).

perspectivă.<sup>210</sup> Conștiința este și ea mai degrabă un aspect al activității omului.<sup>211</sup>

Neisser creionează un contrast între vederile sale din 1967 și cele din 1976. În 1967 accentul cade pe urmărirea "fluxului" informațiilor în sistem (în creier); în 1976 el crede că vechea perspectivă este vulnerabilă din unghiul *validității ei ecologice*.

Cotitura intelectuală a lui Neisser amintește izbitor de aceea a lui Terry Winograd, cu accentul ei pe unitatea dintre limbaj și acțiune și pe rolul conversației (al interacțiunii dintre membrii unui colectiv).

Psihologia cognitivă de factură cognitivistă a fost printre vârfurile de lance ale științei cogniției, inclusiv la noi în țară.<sup>212</sup> În continuare mă voi referi doar la ea.

Psihologia cognitivă are, în linii mari, două variante. Una stă sub semnul paradigmei simbolice, a computaționalismului cu alte cuvinte. Alta stă sub semnul paradigmei subsimbolice, altfel spus a conexionismului. Cum unele dintre capacitățile cognitive par a se preta mai bine abordării simbolice, iar altele celei subsimbolice pare cât se poate de rezonabilă ipoteza că sistemele cognitive ale minții umane sunt de fapt, hibride.<sup>213</sup> O astfel de arhitectură cognitivă hibridă propune, la noi în țară, Mircea Micălea. În varianta lui Micălea, stimulii intră mai întâi în câmpul atenției, pentru ca, pe parcurs, să treacă și prin rețele conexioniste, dar să fie prelucrați și în sisteme cu reguli în sens computaționalist.<sup>214</sup>

Din punct de vedere filosofic, un aspect foarte interesant al psihologiei cognitive îl constituie raporturile acesteia cu behaviorismul. Dacă între computaționalism și behaviorism

<sup>210</sup> *Ibidem*, p.104.

<sup>211</sup> *Ibidem*, p. 122.

<sup>212</sup> Vezi contribuția școlii de la Cluj-Napoca și cartea lui Mircea Micălea, *Psihologia cognitivă*, Cluj-Napoca, Gloria, 1994, 488 p.

<sup>213</sup> Am văzut că există, de altfel, tendința de a construi și sisteme artificiale hibride.

<sup>214</sup> Micălea, *Psihologia cognitivă*, pp. 458-461.



incompatibilitatea este limpede, în schimb raporturile dintre behaviorism și conexionism nu sunt tot atât de clare. Spre deosebire de behaviorism, conexionismul admite reprezentările. Într-unele dintre versiuni, reprezentările sunt locale (ele sunt atribuite unor unități, care pot să fie activate sau nu de alte părți ale rețelei), într-altele ele sunt distribuite (sunt atribuite unor configurații de unități din rețea, o unitate putând participa la mai multe astfel de configurații). În schimb, cele două orientări împărtășesc o presuposiție *asociaționistă*. Rețelele sunt antrenate pentru a obține o legătură, o asociere între stimuli și reacții.

Computaționalismul este nativist; presupune că e nevoie de structuri mentale înnăscute pentru a putea prelucra stimuli veniți din exterior. Antrenamentul, presiunea mediului nu pot rezolva problema. Din punct de vedere teoretic nu ne putem mulțumi cu lanțuri de asociații între stimuli și reacții.

## Testul crezămintelor false

Pot însă filosofii să treacă dincolo de genul de analiză a presuposițiilor și să intre mai direct în jocul dezbaterilor din psihologie?

*Crezămintele* (*beliefs*, în limba engleză) sunt una dintre “bijuteriile coroanei” din teoria cunoașterii<sup>215</sup> și din logica filosofică. Într-o perspectivă mai largă, crezămintele trebuie plasate în perspectiva discuției despre *atitudinile propoziționale*.<sup>216</sup> O atitudine propozițională are forma “*A* că *p*”, unde *A* poate să fie, de pildă, “cred” sau “doresc”, iar *p* este o parte propozițională (propoziție în sens gramatical sau un *conținut propozițional*)<sup>217</sup>. Curentă în logica filosofică este

<sup>215</sup> Pentru cunoaștere și crezăminte vezi A. Phillips Griffiths (ed.), *Knowledge and Belief*, Oxford University Press, 1967.

<sup>216</sup> Vezi Nathan Salmon și Scott Soames (editori), *Propositions and Attitudes*. Oxford University Press, 1988, 282 p.

<sup>217</sup> Ca în “cred că *Pământul este rotund*”. Uneori forma lingvistică firească în limba română poate să ceară altceva în locul lui “că” (precum în “*doresc să merg la teatru*”).

versiunea în care partea propozițională este un *conținut propozițional*. Evident, persoana care are atitudinea poate să varieze. Important este că ea are o *relație* cu un conținut propozițional. În cazul crezămintelor, persoana dă *crezare* conținutului propozițional.

Crezămintele sunt propoziții în sens logic: pot fi adevărate sau false. De pildă, cineva poate să creadă că Budapesta este capitala României. Persoana respectivă are, în acest caz, un crezământ fals (dă crezare unui conținut propozițional fals). Dar o persoană poate să dea crezare și unui conținut propozițional de genul "Andrei știe care este capitala României". Dacă Andrei crede că Budapesta este capitala României, atunci persoana respectivă are un *crezământ fals cu privire la ceea ce este în mintea lui Andrei*.

Ni se pare foarte firesc să atribuim crezăminte false altora. Dacă o persoană a ieșit din sală și, între timp, noi am schimbat subiectul discuției, atunci când persoana revine în sală ni se pare normal să o punem în temă. Ne dăm seama că are un crezământ fals despre ceea ce vorbim și încercăm să corectăm acest crezământ. Oamenii maturi au, în mod normal, această capacitate de a *citi mintea celorlalți*. Pot face oare și animalele sau copiii mici același lucru?

Întrebarea de mai sus a fost pusă sub forma "au maimuțele o *teorie a minții*?" de către Premack și Woodruff.<sup>218</sup> Daniel Dennett, unul dintre cei mai cunoscuți filosofi cognitiști, a propus să se testeze capacitatea maimuțelor de a atribui altora crezăminte false, pentru a putea răspunde la întrebarea dacă au sau nu o *teorie a minții*.<sup>219</sup>

Înainte de a spune care au fost rezultatele testelor empirice, se impune un scurt comentariu cu privire la întrebarea lui Dennett. Ea are la bază o distincție filosofică între realitate și ceea ce credem noi despre ea, de fapt între realitate și minte. A

<sup>218</sup> D. Premack și G. Woodruff "Does the Chimpanzee Have a Theory of Mind?", *Behavioral and Brain Sciences* 4, 1978, pp. 515-26.

<sup>219</sup> Daniel C. Dennett, "Beliefs about Beliefs", *Behavioral and Brain Sciences* 1, 1978, pp. 568-70.

atribui altcuiva doar crezăminte care corespund realității înseamnă implicit a trata persoana respectivă ca pe un simplu depozit de adevăruri. O minte comite și erori. De aceea atribuirea de crezăminte false presupune implicit acceptarea ideii că și cealaltă parte are o minte.

Psihologii au preluat întrebarea filosofică și, mai întâi, s-au făcut experimente cu maimuțe. Premack a raportat rezultate, în linii mari, negative. A venit însă și ideea revoluționară a lui Wimmer și Perner de a aplica testul crezămintelor false copiilor mici.<sup>220</sup> Rezultatele au fost spectaculoase.

Cum au procedat psihologii? Scenariul, desigur simplificat, al experimentului este următorul: copilul urmărește o scenetă în care, să zicem, sunt două personaje, pe nume Ana și Maria. Ana are o ciocolată pe care o pune în sertar și iese din cameră. Cât timp Ana lipsește, Maria mută ciocolata în dulap. Ana revine în cameră. Copilul este întrebat unde crede Ana că se află ciocolata.

Copiii foarte mici spun că Ana crede că ciocolata se află în dulap. Ei se ghidează deci după realitate. Copiii mai mari răspund că Ana crede că ciocolata se află în sertar. Experimentele acestea au fost repetate de mai multe ori și vârsta la care este trecut testul a variat.<sup>221</sup> În mod normal, el este trecut la vârsta de patru ani.

Experimentele consacrate crezămintelor despre crezăminte marchează un moment crucial. O problemă filosofică sofisticată ajunge să intereseze și psihologia. Asemenea preocupări aveau să aducă pe masa de lucru a psihologilor chestiuni filosofice dificile privitoare la logica atitudinilor propoziționale, teoria înțelesului și teoria cunoaște-

---

<sup>220</sup> H. Wimmer și J. Perner, "Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception", *Cognition* 13, 1983, pp.103-28.

<sup>221</sup> Pentru o prezentare succintă a acestor experimente și a cadrului mai general vezi Alan M. Leslie, "Child's Theory of Mind", în *Routledge Encyclopedia of Philosophy*.

rii. Filosofia, la rândul ei, a fost puternic marcată de experimente empirice, fără ca prin aceasta să “piardă subiectul”, ca să spun așa, cum se întâmplă în cazul fizicii. Nu se pune problema ca, de acum înainte, acest subiect să fie doar apanajul unei științe, iar filosofia să analizeze rezultatele obținute. Mai mult decât atât, crezămintele interesează în mod inevitabil și inteligența artificială.<sup>222</sup> Avem toate motivele să tragem concluzia că o chestiune de filosofia minții a fost ferm ancorată în orizontul științei cogniției.

## Interpretarea minților

Ne este familiară ideea de *interpretare* a unui text. Uneori, cu un termen mai pretențios, o denumim *hermeneutică*. De asemenea, ne ciocnim de interpretare, în alt sens decât acela de “hermeneutică”, în logică, mai precis în semantica logică. Într-o carte recentă,<sup>223</sup> Radu J. Bogdan<sup>224</sup> propune să numim interpretare competența care le permite primatelor să atribuie atitudini, intenții, sentimente minților. Cu alte cuvinte, interpretarea este competența care permite “citirea” altor minți, dezvoltarea unei întregi *psihologii populare*.

Sintagma “psihologie populară” se referă la capacitatea minților umane de a se raporta spontan la alte minți în termeni precum “intenții”, “dorințe” ș.a.m.d. și de a produce în acest cadru predicții și explicații ale comportamentului celorlalți. După cum am văzut în capitolul al doilea, behavioriștii au tins să

<sup>222</sup> Pentru utilizarea crezămintelor în programele de inteligență artificială vezi Dean, Allen și Aloimonos, *Artificial Intelligence*, pp. 273-277. Pentru un program de inteligență artificială care folosește crezămintele agentului (artificial) vezi “Oscar”, scris de John L. Pollock (Departamentul de filosofie, Universitatea din Arizona). Programul (sursele în Lisp) și documentația aferentă sunt disponibile pe Internet (v., inclusiv pentru copyright, pagina de web a autorului și trimerile făcute în această pagină <http://www.u.arizona.edu/~pollock>).

<sup>223</sup> Radu J. Bogdan, *Interpreting Minds*, MIT Press, 1997. 304 p.

<sup>224</sup> Cititoarele și cititorii din România sunt familiarizați cu contribuții ale lui Radu J. Bogdan la știința cogniției din *Grounds for Cognition* (1994), tradusă în limba română, la editura ALL, de către Mircea Dumitru.

înlocuiescă toată această teorie naivă și limbajul ei cu ceea ce ei credeau că reprezintă o *psihologie științifică*. Dacă este să ne limităm la mintea umană, interpretarea minților nu este deci altceva decât psihologie populară sau naivă.

În *Temeiuri ale cogniției* Radu J. Bogdan pornește, ca să spunem așa, de la anul zero al cogniției, de la cristale care se pot reproduce. Evoluția precede viața biologică. Selecția naturală operează și în cazul cristalelor. Atâta doar că acestea sunt lipsite de un *scop*. Mai departe, forța argumentului nu se bazează doar pe ideea că selecția naturală, în cazul vieții, va conduce până la urmă la minți, ci pe ideea că selecția este suficientă pentru a explica apariția unor *organisme care au scopuri*. Partea fascinantă a argumentului rezidă ulterior în ideea că aceste organisme care au scopuri au resurse modeste de atingere a lor. În comparație cu forțele potențial distructive care le înconjoară în mediu. De unde rezultă, până la urmă, succesul lor? Din orientarea către un scop. Iar un organism are *cunoaștere* dacă are acces la datele care-i permit să atingă scopul. Organismele pot dirija lanțuri cauzale în așa fel încât să poată atinge scopurile. Iar o minte recurge pentru aceasta la reprezentări.

Nu ne putem noi oare forma crezăminte și despre alte minți? Ba da. Și în felul acesta putem dirija și un fel aparte de lanțuri cauzale! Este vorba de lanțurile care își au temeiul în eficacitatea cauzală a minților.

*Interpretarea minților* este deslușită de Radu J. Bogdan din perspectivă evoluționistă. Așa cum arată și Radu J. Bogdan, deși figuri de marcă ale științei cogniției, precum Daniel Dennett,<sup>225</sup> au evidențiat virtuțile recursului la evoluție, cultivarea valențelor explicative ale acesteia nu este, de regulă, agreată în știința clasică a cogniției.<sup>226</sup>

Din perspectiva orizontului evoluției, interpretarea este o *adaptare*.<sup>227</sup> Ea conferă, cu alte cuvinte, avantaje în procesul de

<sup>225</sup> Despre spiritul evoluționist al lui Dennett vezi Radu J. Bogdan, *Interpreting Minds*, p. 262.

<sup>226</sup> Radu J. Bogdan, *Interpreting Minds*, p. 49.

<sup>227</sup> Radu J. Bogdan, *Interpreting Minds*, p. 25.

selecție. Care ar fi *rațiunile* pentru care ființe precum maimuțele sau oamenii recurg la interpretare? Bogdan distinge trei serii de asemenea rațiuni, pe care el le-a botezat *epistemice, comunitare și politice*.<sup>228</sup> Denumirile acestea sunt sugestive, dar pentru a le prinde înțelesul trebuie urmărite detaliile tehnice. Activitățile comunitare asigură, de pildă, menținerea integrității grupului (luat într-un sens mai general decât cel de grup uman; poate fi și un grup de maimuțe), prin cooperare împotriva prădătorilor, fărâșire de unelte etc. "Politic" are, de asemenea, un sens mult mai larg decât cel uzual. Este vorba despre "competiție, exploatare, vicleșug și alianțe tactice" în grupuri foarte variate. Există politică, în acest sens, și în familie; există politică la mai toate mamiferele care trăiesc în grupuri.<sup>229</sup> În rezumat, cele trei serii de rațiuni corespund unor serii de *presiuni evolutive* generate în procesele de educație și comunicare (presiuni epistemice), viața în grup (presiuni comunitare) și manipulări și vicleșuguri (presiuni politice).

În viziunea lui Radu J. Bogdan, interpretarea ia startul (la copiii mici, de pildă) ca *interpretare situată* (dependentă de un context, de o situație concretă).<sup>230</sup> Ea este o *practică* fasonată de evoluție. De ce? Pentru că la fiecare cotitură a procesului evolutiv, conferă avantaje celor care o practică mai bine.<sup>231</sup> Dacă animalul poate urmări *linia privirii* altei ființe, atunci va putea deduce către ce se îndreaptă atenția celeilalte ființe și care sunt *scopurile* acesteia în lume.<sup>232</sup>

De ce începe, la trei-patru ani, copilul să treacă teste precum cel al crezământului fals? Interpretarea din stadiile inițiale este situată, dependentă de contextul concret. Copilul se îndreaptă însă, în stadiile mai târzii, către o interpretare nesituată. Nu mai depinde doar de ceea ce este aici și acum. Nu vede doar lucruri de genul privirii care se îndreaptă *acum* către ceva. Compară situații. Evaluează situații care nu mai există în

<sup>228</sup> Radu J. Bogdan, *Interpreting Minds*, p. 31

<sup>229</sup> Radu J. Bogdan, *Interpreting Minds*, pp. 31-32.

<sup>230</sup> Cf. Radu J. Bogdan, *Interpreting Minds*, p. 69 ș.u.

<sup>231</sup> Cf. Radu J. Bogdan, *Interpreting Minds*, p. 78.

<sup>232</sup> Vezi analizorul privirii în Radu J. Bogdan, *Interpreting Minds*, p. 138.

momentul actual. Copilul se îndreaptă către o *revoluție mentală*.<sup>233</sup> În contrast cu mainutele, care rămân la o interpretare situată, copiii trec la interpretarea nesituată. Revoluția mentală eliberează de tirania contextului; cultura este cea care reintroduce constrângeri.<sup>234</sup>

Partea cea mai frapantă a argumentului de mai sus este legată de faptul că practici precum cea a recursului la vicleșug sunt strâns legate de fibra intimă a minții umane. Vicleșugul, pentru a fi eficient, presupune o interpretare corectă a intențiilor celuilalt, cuplată cu interpretarea adecvată a crezămintelor și dorințelor, precum în povestea Calului Troian. Ulise interpretează bine mințile troienilor și astfel grecii pătrund în cetate. În felul acesta, interpretarea permite dirijarea lanțurilor cauzale declanșate de alte minți – ceea ce și era problema centrală analizată în acest paragraf.

### Incapacitatea de a interpreta mințile

Cineva poate să aibă handicapul de a nu vedea sau de a nu auzi. S-ar putea ca handicapul să fie strict fizic, precum lipsa unor degete. Relativ recent, în 1943, doctorul Leo Kanner a descris un handicap care ține de dezvoltarea minții. El se concretizează, într-o formă a sa, prin incapacitate a comunica, printr-o rădusă integrare socială. Kanner a denumit acest fenomen *autism*.<sup>235</sup> Alte forme de autism se caracterizează prin combinarea lipsei de interacțiune socială cu o stăpânire normală a limbajului și trecerea testelor de inteligență la nivel mediu și peste medie (variantea Asperger) sau prin pierderea deprinderilor dobândite între unu și patru ani și pierderea deprinderii de a folosi mâinile (variantea Rett), prin dereglarea semnificativă a unor comportamente (autism atipic).

<sup>233</sup> Radu J. Bogdan, *Interpreting Minds*, p. 177.

<sup>234</sup> Radu J. Bogdan, *Interpreting Minds*, p. 200.

<sup>235</sup> Cf. "What is Autism?", pagină de web a Autism Society of America (<http://www.autism-society.org/autism.html>) [actualizare din data de 17.04.1999].

Autismul nu este o boală. Dacă am folosi limbajul lui Chomsky, am putea spune că autiștilor le lipsește un *organ mental*, mai precis le lipsesc capacități legate de interpretarea minților. Rezultatul este că, în unele cazuri, au capacități de ordin logic perfect dezvoltate, dar nu pot pricepe glumele sau nu reacționează normal la tachinare. De asemenea, au serioase dificultăți de a înțelege durerea sau suferința altora.

Cazurile de autism au devenit adevărate teste pentru susțineri filosofice-cheie. Cel mai important pare a fi felul în care autismul pune sub semnul întrebării ideea filosofică, larg răspândită, conform căreia capacitățile logice, cu alte cuvinte *rațiunea*, înțeleasă ca o facultate de a judeca logic sunt centrale pentru mintea umană. Or, pot exista autiști care au această capacitate logică perfect dezvoltată, dar nu sunt deloc buni interpreți ai altor minți, nu pot lega relații de prietenie sau folosi cu ușurință limbajul relațiilor interumane.<sup>236</sup>

În literatura de specialitate, fapte de acest gen au condus, firește, la ipoteza că există un *modul* (v. anterior ideea de modularitate a minții) care este răspunzător de interpretare. După unii autori, acest modul produce *teorii ale minții*. Altfel spus, modulul generează ipoteze și strânge dovezi în legătură cu mintea altora. Dacă modulul este defect, persoana nu “vede” mințile din jur.<sup>237</sup>

Ipoteza că autiștii nu “văd” mințile altora este coroborată de rezultatele obținute de către ei la testul crezămintelor false.

Există însă și o ipoteză aflată în competiție cu ipoteza defectării modulului dedicat teoriei minții. Conform acestei ipoteze, deficiența de care suferă autiștii este mai profundă. Ea

<sup>236</sup> Într-una din prelegerile sale la Universitatea din București, Radu J. Bogdan îi compara cu programele de calculator centrate pe logică, care pot judeca fabulos de repede dar au un aer de robot lipsit de umanitate. Vezi Radu J. Bogdan, *Interpreting Minds*, p. 166, unde scrie că “a avea concepte, drapate în haine lingvistice și procesate logic, nu înseamnă și a avea o interpretare a crezămintelor și intențiilor”.

<sup>237</sup> Cf. Peter K. Carruthers, “Autism as mind-blindness: an elaboration and partial defence”, in P. Carruthers & P. K. Smith (editori), *Theories of theories of mind*, Cambridge University Press, 1966, pp. 257-273.



ține de lipsa unei capacități de a se angaja în exerciții de imaginație, în supoziții (de genul "să zicem că ..."). Această incapacitate îi face pe autiști să nu poată *simula* în mintea lor ceea ce se petrece în mintea altora.

Asemenea ipoteze conduc evident la iluminarea unor aspecte profunde ale funcționării minții umane și la controversele pe care sunt de natură să le stârnească asemenea investigații.

Care ar fi unele argumente în favoarea modulului dedicat teoriei minții? Dacă acesta ar lipsi, atunci persoana respectivă este văduvită de arsenalul de concepte și de capacitatea de a face inferențe cu privire la orice eu. Se pare că acest lucru este susținut de constatările empirice, conform cărora copiii autiști nu sunt în stare să-și amintească nici propriile lor crezăminte false.<sup>238</sup> Alte teste experimentale sugerează incapacitatea autiștilor tip Asperger de a relata trăiri interne (limbaj interior, de pildă).<sup>239</sup>

Pentru ipoteza lipsei capacității de a simula există însă, de asemenea, argumente. Autiștii nu se angajează în jocuri în care copiii pretind că sunt altceva sau altcineva. De aici ideea că ei ar avea, de fapt, o deficiență de imaginație.<sup>240</sup>

## Deschideri către psihanaliză?

Până acum a fost vorba despre capacitatea de a interpreta sau nu și mai puțin despre ascunderea de interpretarea altora. Desigur, în cazul vicleșugurilor, o parte a artei este aceea de a nu permite celorlalți să ghicească intențiile celei sau celui care întinde o capcană.

<sup>238</sup> Carruthers, *Autism*, p. 260.

<sup>239</sup> Carruthers, *Autism*, p. 261.

<sup>240</sup> Carruthers susține că teoria conform căreia autiștii nu "văd" alte minți poate explica aceste cazuri prin lipsa de *motivație*: copiii autiști nu au acces la propriile trăiri și nu pot *gusta plăcerea* de a-ți închipui că ești o altă persoană sau că ești în altă parte.

Clasică este în psihanaliză explicația lapsusurilor. Acestea ar revela prezența unor dorințe ascunse. Ascunse chiar față de subiect. Ca și în cognitivism, în psihanaliză structura *mentală internă* este bogat articulată și prezintă o capacitate de a trece dincolo de lanțuri de stimuli și reacții. Specifică însă psihanalizei este ideea că sistemul minții are o structură *multinivelară*.<sup>241</sup> Are cel puțin două niveluri, unul la care propriile atitudini propoziționale sunt accesibile subiectului și altul la care acestea nu sunt accesibile, dar au o eficacitate cauzală.

Din această perspectivă, cel de al doilea nivel, cel ascuns privirii interpretative, găzduiește *dorințele inconștiente*. Acestea au acces, din când în când, la căile care conduc către ieșirile din sistem. Dacă ar fi să adoptăm anumite versiuni ale psihanalizei, de pildă cea care-și are originea în opera lui Lacan, mecanisme de tipul celor ale unui limbaj sunt în funcțiune la nivelul inconștientului.

Dacă încercăm să ne apropiem însă de versiunea clasică a psihanalizei, ceea ce vedem pare a fi mai degrabă un mod energetist de a înțelege mintea. Inconștientul “fierbe” și “supapele” lasă, din când în când, să scape vaporii fierbinți. Un asemenea model se apropie mai curând de mecanicismul atracției și respingerii ideilor despre care vom vorbi în continuare.

### Trei modele ale minții

Acest paragraf încearcă să tragă o linie sub care să așeze o concluzie cu privire la drumul parcurs până acum în filosofia minții. De ce acum și nu după capitolul următor? Cred că motivele vor deveni evidente fie și din titlul respectivului capitol. Oricum, vor exista și acolo două modele ale minții și

---

<sup>241</sup> Pentru raporturile între știința cognției și psihanaliză vezi Évelyne Caralp și Max Gallo, *Sciences cognitives et psychanalyse*, Toulouse, Milan, 1998, pp. 44-49.

cititoarele sau cititorii sunt liberi să încerce să le combine cu modelele de mai jos.

Ar mai fi de adăugat încă o observație preliminară. Titlul acestei paranteze rezumative este "Trei modele" și nu "Cele trei modele". Nu există deci pretenția exhaustivității. Tot ce aş susține doar este că nu există multe modele ale minții. Într-un fel, acest lucru este surprinzător. El arată, cred, că modelele minții corespund unor mari viziuni din istoria științei și a tehnologiei, precum și unor diviziuni în câmpul cercetărilor: o viziune mecanică și energetistă, una centrată pe științele și disciplinele umaniste și o viziune informațională. Corespondența nu este strict biunivocă, dar analogiile sunt frapante.

- **Modelul asociaționist.** Termenul "asociaționist" este teribil de simplificator, dar el sugerează ceva: mintea face legături între ceea ce se petrece în jurul ei și comportamente. Behavioriștii au considerat că secretele minții pot fi surprinse dacă studiem meticulos lanțuri lungi de asemenea legături. Pentru aceasta nici nu este nevoie să postulăm entități mentale. Legătura este totul. În alte versiuni ale modelului accentul este mai puternic pe atracția și respingerea ideilor. Iar în alte variante pe ceva care seamănă cu transformările energetice. Obiectele nu sunt oglindite de către minte, ci ele se transformă în energii ale dorințelor noastre, pe care le refulăm și le defulăm.
- **Modelul acțional.** Din nou termenul "acțional" poate să pară prea restrâns. Ideea modelului este legată însă strâns de imposibilitatea de a separa mintea de acțiune. Omul este o ființă implicată în lume. Este o ființă activă. Nu obiectele vin peste om, ci omul selectează din mediu. În acest model unitatea dintre minte și lume, precum și interacțiunile dintre minți formează întreguri care nu pot fi înțelese dacă sunt disecate.
- **Modelul computațional.** În acest model mintea are o structură internă bogată, iar cunoașterea de care dispune mintea sub diferite forme este elementul care contribuie decisiv la rezolvarea problemelor cu care ea se confruntă. Structura internă a minții are o arhitectură

compusă din dispozitive capabile să preia, să memoreze, să prelucereze date și să dirijeze fluxuri de date și procese de prelucrare a datelor, precum și să le transmită către ieșirile în exterior. Această arhitectură are o parte fizică și o parte virtuală. Proprietatea crucială a acestei arhitecturi este marea ei flexibilitate. Pot exista diverse metode de exploatare a flexibilității: schimbarea părții virtuale prin programarea dispozitivelor fizice; reconfigurarea dispozitivelor prin schimbarea unor parametri la nivelul structurilor subiacente dispozitivelor. Poate exista o prelucrare centralizată a datelor sau una în paralel.

## Misterele minții

Dacă nu ar fi cam ciudat ca numai acest capitol să aibă motto, atunci sigur că extrasul care ar servi cel mai bine acestui scop ar fi “conștiința este cel mai mare mister”<sup>242</sup>.

Într-un fel așa este: *misterele* au în nucleul lor un mister central, de la care se trag toate celelalte. Cum se întâmplă de multe ori, trecerile sunt însă graduale. Nu tot ce nu ține de conștiință este limpede și nu tot ce ține de conștiință este misterios.

Capitolul începe cu câteva date relativ tehnice despre reprezentarea cunoașterii din perspectiva sistemelor artificiale. De ce ar fi notele tehnice despre reprezentarea cunoașterii incluse aici? Evident, nu este nimic misterios în aceste tehnici dacă le luăm doar ca simple tehnici. Dacă ambiția este însă aceea de a *reprezenta cunoașterea* așa cum o reprezintă și mintea, atunci semnele de întrebare sunt numeroase. Ele vor fi puse în evidență de distincția lui Chalmers între două modele ale minții. Abia apoi urmează o trecere în revistă a unor versiuni ale abordării filosofice a conștiinței.

Capitolul nu are rezumate sau concluzii. Prezentarea cărții lui Chalmers despre conștiință se încheie însă cu concluzia că există temeiuri pentru a crede că studiul sistemelor computaționale conduce la o lărgire mult mai substanțială și surprinzătoare a orizontului nostru decât ni se pare la prima vedere. Un paragraf cu obiectivele atinse în acest sens ni se pare însă nepotrivit în cazul unui domeniu dinamic, cum este cel constituit de filosofia minții și știința cogniției. Ar însemna să anunțăm că acest ultim mare mister care este mintea a fost pătruns, ceea ce nu este cazul.

Oricum, acesta nu este ultimul capitol, deoarece bibliografia este aranjată sub forma unui ghid, care poate fi citit

---

<sup>242</sup> David Chalmers, *The Conscious Mind. In Search of a Fundamental Theory*, Oxford University Press, p. Xi.

și ca un rezumat al rezultatelor obținute și al problemelor ce-și așteaptă rezolvarea.

## Câteva repere ale reprezentării cunoașterii

De multe ori, în textul de mai sus am făcut apel la ideea de *reprezentare*. Lucrurile au fost tratate însă din punctul de vedere care mi se pare prudent: reprezentările sunt prelucrate computațional; deci ceea ce intră și ceea ce iese din procesele computaționale constituie, dacă vrem să gândim așa, "reprezentări". Altfel, și mai prudent pare să vorbim despre *date*. Iar cel mai prudent este să vorbim, abstract, despre listele de biți, cum am și făcut-o adesea.

Istoria reprezentării cunoașterii în sisteme artificiale este cât se poate de recentă. Ea începe practic cu primii pași ai inteligenței artificiale. Iar a vorbi despre reprezentări în aceste sisteme înseamnă în primul rând o referire la structuri de date adecvate reprezentării cunoașterii.

După cum am arătat, sistemele expert, cel puțin din punct de vedere comercial, formează partea cea mai reușită a sistemelor computaționale artificiale. Nu se presupune că un sistem expert *are cunoaștere* în sensul în care spunem că oamenii posedă cunoaștere. Sistemul are însă o reprezentare a cunoașterii (de care dispune expertul uman). Această reprezentare în programul de computer are forma extrem de simplă a unor *reguli*. După cum se poate vedea din exemple prezentate aici, limbaje dedicate sistemelor expert, precum CLIPS, folosesc o notație de genul:

FAPTE

=>

ACȚIUNI.

Dacă sistemul găsește în baza sa de date FAPTE-le, atunci execută ACȚIUNI-le. Majusculele sunt folosite aici

intenționat, deoarece "FAPTE" înseamnă doar șiruri de simboluri formale corespunzătoare unor stări de lucruri, iar "ACȚIUNI"-le sunt operații cu simboluri formale.

Din punctul de vedere al filosofilor, o reprezentare familiară a cunoașterii este *rețeaua propozițională*.<sup>243</sup> Abstracția, în sensul tehnic din programare, a unei rețele semantice este o mulțime de *noduri* (sau puncte) și de *legături* între noduri. Legăturile pot fi de diverse tipuri (marcate prin etichete corespunzătoare). O propoziție este reprezentată sub forma unor noduri între care există legături. De exemplu, "Socrate" și "om" pot fi două noduri. Dacă nu au legătură, nu există o propoziție. Dacă legătura este de tipul "este", atunci propoziția e "Socrate este om".

Din punctul de vedere al elaborării de programe de computer, reprezentările de mai sus sunt foarte importate pentru că definesc structuri de date. O dată definite aceste structuri este posibil să fie proiectate metode de prelucrare a datelor care vor fi incluse în structurile respective. Iar dacă structurile de date și metodele sunt "împachetate" într-un singur tot, se poate construi un obiect. Cum am văzut deja, din obiecte se pot făuri programe.

De exemplu, pe o rețea propozițională este posibil să fie definită o metodă de generalizare existențială: de la "Socrate este om" se poate trece la "Există un om".

Un alt mod de a reprezenta cunoașterea este cel al *cadrelor* (în engleză "frame"). Am putea sugera felul în care arată abstracția unui cadru comparându-l cu un formular de cerere. Formularul necompletat are *rubrici* în care se pun date de bază sau date structurate complex (inclusiv sub forma unui cadru).<sup>244</sup>

<sup>243</sup> Denumită și *rețea semantică*. Vezi Giarratano and Riley, *Expert Systems*, pp.69 ș.u. pentru o expunere introductivă.

<sup>244</sup> Pentru o prezentare detaliată și exemple vezi H.Ueno et.al., *The Representation and Use of Knowledge*, cap. 3 (originalul este în Japoneză; am folosit traducerea în rusă: Moscova, Mir, 1989).

Visul lui Turing a fost acela că într-o zi un program va putea să se comporte ca o minte umană în asemenea măsură încât să nu poată fi distins de aceasta decât prin suportul fizic.

O parte cheie a visului lui Turing pare să fie însă legată de *înțelegerea* de către program a reprezentărilor. Dacă eu am o reprezentare a unei propoziții despre pisica noastră și o folosesc conștient, atunci sunt capabil să *înțeleg conținutul propozițional* aferent. Mai simplu, orice reprezentare aș avea despre motanul de acasă, dacă o folosesc conștient, atunci *înțeleg* că este vorba despre motanul de acasă.

După cum am văzut deja, adversarii inteligenței computaționale au folosit observații precum cea de mai sus pentru a respinge în bloc întregul efort de a construi sisteme artificiale inteligente. Pentru Dreyfus această încercare este o alchimie a secolului al XX-lea. Distincția care pare necesară este însă alta. Este o distincție între nivelurile (sau părțile) minții care funcționează automat și cele care funcționează conștient. Paradoxal, știm să le modelăm mai bine pe primele, deși avem o cunoaștere directă doar a celor din urmă.

Atunci când vorbim despre părțile accesibile și inaccesibile ale minții ar trebui amintit că filosofii își fac mari probleme din accesul la mintea *altora*, ba consideră chiar că este o mare problemă demonstrarea existenței *altor minți*. Cum s-ar spune, chestiunea arzătoare este vindecarea autismului filosofic.

Să presupunem însă, pentru o clipă, că Robinson a ajuns pe insula sa și suferă de o amnezie completă: orice este legat de cultura umană a fost uitat. El poate însă continua să acționeze rațional, în sensul că este un maximizator de utilități. Robinson supraviețuiește. Teoria alegerii raționale ne ajută să înțelegem cum se poate petrece acest lucru.<sup>245</sup>

---

<sup>245</sup> Vezi David Friedman, *Price Theory*, ediția a doua, Cincinnati, South-Western Publishing Company, 1990, pp.2-15 (întreaga carte este o excelentă introducere în abordarea din perspectiva alegerii raționale a comportamentului uman).



Cheia experimentului imaginar de mai sus este legată de ideea că acțiunea rațională fără minte este perfect posibilă.<sup>246</sup> Ce poate face Robinson poate face și un computer programat în mod adecvat. Mintea noastră este problema reală. Există ea oare? Mai precis, s-ar părea că problema celorlalte minți nu poate fi despărțită de problema propriei noastre minți.

După cum am văzut, cercetările recente din știința cogniției au scos la iveală importanța capacității minții umane de a recunoaște existența altor minți și de a le interpreta. O parte a soluției problemei de mai sus se spune că ar sta în recunoașterile reciproce. Ce se face însă Robinson, cel puțin cât timp este singur pe insulă?

Pentru a specula în legătură cu un posibil răspuns, aș face o paranteză și aș menționa caracterul aparent superfluu al unora dintre distincțiile din limbile naturale. Sunt limbi care au anumite categorii. Altele nu le au, dar funcționează cât se poate de bine și fără ele. Iar, în sfârșit, limbajele logice tind evident să nu se încarce cu prea multe categorii.

Un exemplu frapant în acest sens îl constituie *genurile*. De ce există genuri? O posibilă sursă a lor este în credințele religioase cu privire la obiecte și fenomene.<sup>247</sup> Nu pare să existe un temei logic pentru genuri.

Ce ar putea readuce mintea în experimentul imaginar cu Robinson? Să zicem că Robinson își închipuie că tunetul este o voce care-i transmite un mesaj. Presupoziția este, în acest caz, din perspectiva lui Robinson, că există ceva care prelucrează cunoaștere și are o reprezentare a existenței lui Robinson în lume (căci îi transmite un mesaj). Jocul în care se angajează Robinson îl pune în aceeași situație în care sunt oamenii care recunosc existența altor minți.

---

<sup>246</sup> Pentru ideea "raționalității fără minte" vezi discuția din David Friedman, *Price Theory*, pp.15-17. El se referă explicit și la programele de computer care pot învăța.

<sup>247</sup> Cf. Götz Wienold, *Genus und Semantik*, Meisenheim am Glan, Verlag Anton Hain, 1967, pp.110 ff. Vezi întreaga carte pentru problema genurilor în lingvistică.

Ipoteza este jocul în care s-a angajat Robinson va reintroduce mintea în peisajul experimentului imaginat mai sus.

Care ar fi morala? Minteă umană, pe de o parte, nu este cu adevărat o minte dacă nu există și reprezentări *conștiente*. Pe de altă parte, ceea ce numim cultură umană joacă un rol în funcționarea minții. Religia este o parte a acestei culturi. Chiar dacă ar fi doar un produs al imaginației, ea tot este importantă, contrar celor susține de atei și liber-cugetători. Este un mod, de fapt, de a produce un exercițiu de imaginație aparte, care creează o minte care privește *ca pe o minte* pe cel care o imaginează.

## Două concepții despre minte

Unul dintre reperele principale în peisajul ultimilor ani în filosofia minții îl constituie cartea lui David Chalmers *The Conscious Mind*. În ea, Chalmers distinge, ca și alți autori de altfel, două tipuri de concepții despre minte, care corespund în fapt cu două aspecte ale minții. Dacă veți avea răbdarea să parcurgeți prezentarea din paragrafele care urmează, veți vedea însă la ce poziție surprinzătoare ajunge Chalmers.

Știința cogniției, scrie Chalmers, deși nu are aproape nimic de spus despre mintea conștientă, are totuși o mulțime de lucruri de zis despre minte în genere. Știința cogniției furnizează explicații ale comportamentelor și este preocupată de acele stări mentale care sunt cauzal relevante pentru comportament. Găsim aici rădăcina unei *concepții psihologice despre minte*.<sup>248</sup>

Cealaltă concepție posibilă despre minte are la bază mintea văzută ca *trăire conștientă*. Din punctul de vedere al acestei concepții, există stări mentale trăite conștient. Acestea nu epuizează spectrul stărilor mentale, dar constituie un aspect dificil de înțeles al minții.

Cele două concepții despre minte, arată Chalmers, nu se exclud una pe alta. Ele vizează fenomene diferite.<sup>249</sup> Ele sunt

<sup>248</sup> Chalmers, *The Conscious Mind*, p. 11.

<sup>249</sup> *Ibidem*.

menite să explice însă prezența, între altele, a unor diverse moduri de a înțelege conștiința.

Pe de altă parte, după Ned Block, *conștiința* este un concept-cupolă, un fel de *mall* sub acoperișul căruia se vând fel de fel de lucruri.<sup>250</sup> De pildă, Block arată că, pentru a explica fenomene legate de conștiință, trebuie să distingem o *conștiință fenomenală* de o *conștiință care mediază accesul*. Prima are un conținut fenomenal, a doua are un conținut reprezentational.<sup>251</sup>

La rândul său, Chalmers pune aspectele minții în corespondență cu două moduri de a înțelege conștiința, ca o conștiință *psihologică* și ca o conștiință *fenomenală*.<sup>252</sup> Trăsătura caracteristică pentru conștiința psihologică o constituie capacitatea de a raporta (de exemplu, verbal) stări mentale în care se află subiectul.

Nu este nimic misterios, până la urmă, crede Chalmers. În conștiința psihologică, deoarece este limpede că un sistem poate să scaneze stările sale interne și să folosească datele astfel obținute în procesele computaționale. Adevărata problemă o reprezintă conștiința fenomenală.

În orice caz, o primă concluzie este limpede. Termenul “conștiință” trimite la un mănunchi de concepte reunite sub o singură cupolă. Folosirea sa trebuie să includă și precizarea tipului de conștiință vizat.

## Superveniența

Superveniența este un *cadru filosofic* în care pot fi duse nu doar discuții privitoare la filosofia minții. Îl veți găsi, de pildă, în filosofia morală. De asemenea, deși conceptul de *superveniență* a fost creat în secolul al XX-lea, probleme de

<sup>250</sup> Ned Block, “On a confusion about a function of consciousness”, *Behavioral and Brain Sciences* 18, 2/1995, pp. 227-287.

<sup>251</sup> *Ibidem*.

<sup>252</sup> Chalmers, *The Conscious Mind*, pp. 25-26.

genul celor care se pun în cadrul supervenienței se ridică și atunci când se vorbește despre proprietăți emergente.<sup>253</sup>

Abstract vorbind, pentru a avea *cadrul* supervenienței este nevoie de două *mulțimi de fapte*. O mulțime va fi numită *subiacentă* sau de bază, iar cealaltă *supervenientă* sau de nivel superior. Principiul fundamental al cadrului este următorul: dacă este determinată regula apartenenței la mulțimea subiacentă, atunci este fixată și regula apartenenței la mulțimea supervenientă. Implicația inversă nu este adevărată.

Exemplul cel mai la îndemână aici este cel al sistemelor computaționale. Mulțimea faptelor subiacente este cea a proprietăților dispozitivelor fizice. Faptele privitoare la mașinile virtuale aparțin mulțimii superveniente. Pentru ca un program să poată fi rulat pe una dintre mașinile virtuale, dispozitivele fizice trebuie puse în funcțiune. Inversând lucrurile însă, același program poate rula pe sisteme fizic foarte diferite. Mașinile virtuale pot fi construite în sisteme de dispozitive fizic foarte diferite.

Un lucru ar trebui poate precizat. Dacă anumite fapte sunt constatate la nivelul supervenient, atunci nivelul subiacent nu poate arăta în așa fel încât aceste fapte să fie imposibile. Legile fizicii nu pot fi, din acest motiv, acceptabile dacă exclud existența vieții sau a omului (este așa-numitul principiu antropic).

De ce este atrăgător cadrul supervenienței pentru filosofi? Răspunsul este destul de simplu: (1) cele două niveluri nu sunt reduse unul la altul (vezi mai sus exemplul cu programele care nu pot fi reduse la o anume mulțime de evenimente fizice); (2) nu se instalează nici un dualism greu de explicat (programele, pentru a fi stocate sau executate, au nevoie de o mulțime subiacentă de fapte fizice).

Cadrul conturat mai sus este cadrul de bază. Pentru a dezvolta argumente sofisticate, filosofi extind acest cadru.

<sup>253</sup> Simon Blackburn, "Supervenience". *Routledge Encyclopedia of Philosophy*, Version 1.0

Principalul mijloc de extindere a cadrului îl constituie *lumi posibile*. Pentru nevoile prezentării care urmează "lumile" ar putea fi tratate chiar prin analogie cu "lumea" în care trăim. găsim în ea indivizi, proprietăți, relații. Prin abstracție, aceștia pot deveni indivizi, proprietăți și relații în sens logic. Tehnic vorbind, principalul avantaj al unei lumi posibile este că putem relativiza conceptul de *adevăr* al unei propoziții la lumi posibile. Astfel și mulțimile de *fapte* despre care este vorba mai sus sunt relativizate la lumi posibile.

În cadrul astfel îmbogățit, se pot face noi distincții. Chalmers, de pildă, distinge între superveniență *locală* și superveniență *globală*. O mulțime de proprietăți *B* este superveniență local peste o mulțime de proprietăți *A* subiacente dacă orice individ *i* și orice individ *j* care au aceleași proprietăți *A* au și aceleași proprietăți *B*.<sup>254</sup>

Conceptul de superveniență locală este oarecum limitat, deoarece o copie exactă a Columnei lui Traian nu are exact aceeași valoare ca și originalul. Lucrul acesta se vede chiar mai bine atunci când copia este pusă exact *în locul originalului*, cum este cazul leilor din Delphi (duși, pentru a fi protejați, într-un muzeu).

Superveniența este *globală* dacă este definită pentru o întreagă lume posibilă. Dacă două lumi posibile *l* și *w* au aceleași proprietăți subiacente, atunci au și aceleași proprietăți superveniente.

Mai importantă, după Chalmers, este distincția dintre superveniența *logică* și cea *naturală*. Aici intră în joc felul în care este definit *universul accesibil* de lumi posibile printr-o relație de *accesibilitate* *R*: dacă *l* este o lume posibilă, atunci *w* aparține universului accesibil dacă și numai dacă există relația *R* între *l* și *w*. Mai simplu spus, lumea *w* este o alternativă la *l*, dacă și numai dacă este adevărat că *l* are relația *R* cu *w*. Dacă relația *R* înseamnă *logic accesibilă* (posibilă *logic*), atunci universul accesibil este cel al lumilor *logic* posibile. Dacă *R* înseamnă că *w* este o alternativă *naturală* (altfel spus, că *R*

<sup>254</sup> Chalmers, *The Conscious Mind*, pp. 33-34.

înseamnă că în  $l$  și  $w$  sunt valabile aceleași legi ale naturii), atunci universul accesibil este cel al lumilor *natural* posibile.<sup>255</sup>

Se observă că, dacă înțelegem prin *lume posibilă* o situație, atunci *univers accesibil* nu este altceva decât o mulțime de situații alternative.

Spuneam mai sus că *adevărul* este relativizat la lumi posibile. Cu alte cuvinte, fiind date *condiții de adevăr*, ele sunt îndeplinite sau nu de către o lume posibilă din universul dat. Chalmers distinge însă între două mulțimi de condiții de adevăr asociate unui enunț. Pentru aceasta trebuie să examinăm ceva mai îndeaproape ideea de "condiție de adevăr".

Ce înseamnă că enunțul " $x$  este pătrat" e adevărat? Cu toții vom încerca să vedem ce *se înțelege* prin "pătrat". Poate vom descoperi ceva de genul "are toate laturile egale". Are mai puțină importanță cum anume formulăm ideea de pătrat. Important este că, într-o lume  $l$  dată există o mulțime a obiectelor care sunt pătrate. Aceasta este *extensiunea* lui "pătrat" în lumea  $l$ . Dacă luăm însă în considerare felul în care arată extensiunea lui pătrat în orice lume alternativă  $w$ , atunci spunem că am definit *intensiunea*.

Conceptele de mai sus sunt foarte alunecoase de îndată ce nu lucrăm cu termeni precum "pătrat". Am văzut deja acest lucru în discuția despre argumentul lui Kripke împotriva materialismului în filosofia minții.

Filosofilor, spre deosebire de istoricii tradiționaliști, le plac lumile *contrafactice*. Ei și-ar pune bucuroși întrebări de genul "ce se întâmpla dacă Napoleon nu ar fi pierdut bătălia de la Waterloo?" dacă ar putea construi astfel un experiment imaginar care să aibă vreun rost.

Fie  $r$  o lume reală și  $w$  o lume contrafactuală în raport cu lumea  $r$  (o lume în care se întâmplă fapte care nu se întâmplă să fie adevărate în lumea reală, fapte de genul victoriei lui Napoleon la Waterloo). Să zicem, mergând pe urmele unui exemplu celebru, că vrem să fixăm intensiunea termenului

<sup>255</sup> Chalmers, *The Conscious Mind*, pp. 34-38.

“apă”. Chalmers distinge două moduri de a face acest lucru. Fie pornim de la lumea luată drept reală, lumea  $r$ , și în orice lume alternativă extensiunea termenului “apă” este fixată *ca și cum* aceasta ar fi lumea reală, atunci se spune că s-a definit *intensiunea primară*. În cazul concret, în orice lume posibilă alternativă “apă” se referă *în sens primar* la lichidul pe care îl bem, la lichidul din lacuri, mări etc. Dacă se întâmplă ca o lume alternativă să fie *ca și* lumea reală, atunci “apă” se referă la  $H_2O$ . Dacă extensiunea în orice lume posibilă alternativă este fixată *ca și cum* aceasta ar fi o lume contrafactuală, atunci rezultatul este *intensiunea secundară*. În cazul termenului “apă” aceasta ar putea însemna că “apă” se referă în sens secundar la vreo substanță cu o compoziție chimică  $X_nZ$ .

Se vede acum de ce pot exista, în această versiune, două serii de condiții de adevăr. Dacă evaluăm un enunț de genul “ $x$  este apă” din punctul de vedere al intensiunii primare a termenului “apă”, obținem o serie de condiții de adevăr. Dacă folosim intensiunea secundară, obținem cealaltă serie. Similar, vom avea de a face și cu două *conținuturi propoziționale* ale unui enunț.<sup>256</sup>

Date fiind aceste noi distincții și mașinăria supervenienței, dacă folosește lumi posibile, devine mult mai complexă. Care este intenția care stă în spatele acestei mașinării conceptuale? Foarte simplu spus, Chalmers are nevoie de un spațiu logic în care să poată face *analize conceptuale* (să tragă concluzii, cu privire la ce este necesar sau nu, din analiza conceptelor). Intensiunile primare se pretează unui asemenea tip de analiză. Intensiunile secundare țin de zona lui *a posteriori*. În fond, ceea ce se încearcă este punerea cât mai bine la adăpost de obiecții a concluziilor (surprinzătoare !) care vor urma. De aceea aceste distincții însușite de către Chalmers sunt plasate aici și nu în contextul argumentelor lui Kripke contra materialismului în filosofia minții. La un prim contact cu argumentele prezentate în

<sup>256</sup> Vezi Chalmers, *The Conscious Mind*, pp. 56-65 pentru detaliile acestei concepții despre caracterul bidimensional al referentului. Chalmers sintetizează o întreagă literatură produsă de autori precum Putnam, Kripke, Kaplan și alții.

continuare, aceste măsuri de precauție filosofice pot fi însă ignorate.

### **Tipuri de explicații și experimente imaginare cu ființe zombi .**

Superveniența are un rol semnificativ de jucat dacă vrem să distingem între diverse tipuri de explicații. Dacă se explică existența unui fapt de nivel superior (un fapt la nivelul proprietăților superveniente) prin fapte de la un nivel subiacent, atunci *explicația* este una de tip *reductiv*.

Exemplul clasic de explicație reductivă în filosofia minții îl constituie explicațiile propuse de fizicalism, despre care am mai vorbit. În acest caz, nivelul subiacent este cel al proprietăților fizice. Fizicalismul nu este altceva decât un proiect filosofic de a oferi explicații reductive în termenii proprietăților fizice pentru orice fel de proprietate mentală.

O explicație este nonreductivă dacă ea face apel la fapte de același nivel sau la fapte de nivel superior. Explicarea acțiunilor, de pildă, drept rezultat al unor crezăminte și dorințe este o explicație *nonreductivă*. Dacă am lua în considerare efectul fizic ca atare al acțiunii (“piatra care a spart geamul”), atunci este vorba chiar de explicarea unor fapte subiacente, fapte de la nivel fizic, printr-o serie de fapte de nivel superior (“x a crezut că ...” și “x a dorit ca să ...”).<sup>257</sup>

În știința cogniției, explicațiile sunt de tip *funcțional*, ele fac apel la funcționarea unui sistem (abstract) de dispozitive. Se explică felul *cum* o mașină virtuală care face cutare sau cutare lucru poate fi realizată (fie prin programare explicită, fie prin antrenarea unei rețele) din alte dispozitive. Aceasta este o explicație de tip *reductiv*. În principiu, ea ar putea fi transformată într-o explicație care recurge la stări

<sup>257</sup> Chalmers, *The Conscious Mind*, pp. 43 – 48, pentru o discuție mai amănunțită a diverselor genuri de explicații.



neurofiziologice: cutare stare a rețelelor de neuroni reali este răspunzătoare de o anume funcție a sistemului.<sup>258</sup>

Se poate recurge oare la explicații reductive și în cazul conștiinței? Ideea lui Chalmers este că acest lucru este posibil în cazul conștiinței psihologice. Nu este posibil în cazul conștiinței fenomenale. Chalmers trece în revistă mai multe argumente menite să arate acest lucru. Unul dintre aceste argumente este construit în jurul experimentului imaginat de Frank Jackson, pe care l-am prezentat deja.<sup>259</sup> Ce este impresionant însă în cartea lui Chalmers este faptul că este produsă o întreagă serie de argumente care converg toate către o concluzie antifizicalistă.

Cele mai spectaculoase sunt argumentele care folosesc *zomboizi*. Termenul de “zombi” este preluat din literatura științifico-fantastică.<sup>260</sup> Din această perspectivă, o ființă *zombi* este “un corp fără suflet, înviat din morți, și folosit drept sclav”.<sup>261</sup> Aspectul interesant în această definiție este lipsa de suflet. Un zombi are corp, este viu, dar nu are “suflet”.

Din perspectivă filosofică, definiția de mai sus devine “o ființă care se comportă ca și noi și ar putea avea aceeași organizare funcțională ca a noastră, ba chiar aceeași structură neurofiziologică, dar nu are trăiri conștiente sau *qualia*”<sup>262</sup>. Personajul din experimentul lui Jackson era lipsit de trăirea culorilor pentru că stătea închis într-o cameră de unde totul se

<sup>258</sup> Chalmers, *The Conscious Mind*, pp. 46-47, pentru acest mod de a vedea explicațiile din știința cogniției.

<sup>259</sup> Cf. Chalmers, *The Conscious Mind*, pp. 103-104.

<sup>260</sup> Dacă vă uitați pe Internet, veți găsi imediat referiri la povești, filme sau jocuri cu zombi. Un asemenea joc este, de pildă, *Zombie Dinos from Planet Zeltoid*, în care dinozauri zombi sunt pe punctul de a distruge planeta noastră (<http://www.interplay.com/games/zombie.html>).

<sup>261</sup> Definiție apud L. S. Hauser, *Zombies Invade Philosophy*. <http://members.aol.com/lshauser/zomboid.html> [accesat 23.11.99]. Această pagină este foarte utilă pentru că are numeroase legături cu alte pagini despre zombi. Există aici și trimiteri la texte filosofice care conțin experimente imaginare cu zombi.

<sup>262</sup> Larry Hauser, “Zombie”, în *Dictionary of Philosophy of Mind*, pe web <http://www.artsci.wustl.edu/~philos/MindDict/zombie.html> [accesat 23.11.99].

vede în alb și negru. Un zombi nu are însă deloc trăiri. Dar poate funcționa aidoma unei ființe ca noi.

Este oare o ființă zombi posibilă? Dar în ce sens ar fi ea posibilă? Primul argument al lui Chalmers este că ființele zombi (în sens filosofic) sunt logic posibile. Putem examina, de pildă, o ființă zombi care este, din punct de vedere fizic și funcțional, exact ca și mine. Este o copie a mea în altă lume. Are chiar și conștiință psihologică. Nu are o conștiință fenomenală. Este logic posibilă pentru că este conceptual coerentă. Chalmers crede că *onus probandi* cade pe umerii celei sau celui care vede vreo contradicție logică într-un astfel de concept de zombi.

Se pare că un rol decisiv îl joacă aici ideea că un alt sistem poate fi funcțional echivalent cu mine, fără să fie *fenomenologic* echivalent. Mai simplu spus, intuiția este că explicația funcțională lasă loc, logic vorbind, unui sistem care nu are trăiri, dar se comportă aidoma unui om.<sup>263</sup>

S-ar putea ca argumentul de mai sus să nu convingă pe toată lumea. Obiecția ar putea fi de tipul “conștiința fenomenală” joacă totuși un rol în funcționarea minții (oricare ar fi acest rol). Există atunci un alt experiment imaginar, al cărui rezultat este că unor seturi de fapte fizice identice le corespund trăiri conștiente diferite. De data aceasta lumea alternativă este populată cu ființe care sunt fizic identice cu mine sau cu tine și, pe deasupra, sunt conștiente în sens fenomenal. Chalmers arată că este posibil ca trăirile conștiente să fie diferite. Acolo unde eu văd (în sens fenomenal) “roșu” omologul meu vede “verde”. Este așa-numitul *spectru inversat*.<sup>264</sup>

Piesa cheie din angrenajul argumentului pare să fie următoarea: să zicem că omologul (meu) este doar funcțional echivalent cu mine, dar este făcut din siliciu. Atunci chiar că

<sup>263</sup> Nu întâmplător la originea acestor experimente cu zombi se află Searle. Despre Searle și zombi vezi comunicarea lui Larry Hauser, “Revenge of the Zombies” la Colocviul A.P.A. (filiala de Est), 29 decembrie 1995 (text pe web la <http://members.aol.com/lshauser/zombies.html> [accesat 23.11.99]).

<sup>264</sup> Chalmers, *The Conscious Mind*, pp. 99-101.

este conceptual coerentă ideea posibilității logice a inversării spectrului.

Alte argumente sunt legate de asimetria epistemică și absența analizei. Asimetria epistemică este un nume pentru situația (ciudată) de a nu avea probe despre existența conștiinței decât din observațiile noastre interne. Alte observații sunt doar indirecte. Eliminativismul în raport cu conștiința este implauzibil numai pentru că noi suntem familiarizați cu trăirile conștiente. Absența analizei, pe de altă parte, înseamnă lipsa unei analize funcționale a conștiinței fenomenale.

Pe scurt, concluzia seriei de argumente aduse de Chalmers este că faptele legate de conștiința fenomenală nu sunt logic superveniente peste faptele fizice. Desigur, în funcție de rafinarea ideii de superveniență, se pot opera distincțiile de rigoare. Concluzia la care se ajunge ar fi însă tot aceea că nu este posibilă o explicație reductivă a conștiinței fenomenale.

### Conștiință fenomenală și inteligență artificială

După concluzia că faptele ce țin de conștiința fenomenală nu sunt logic superveniente peste fapte fizice, să trecem în revistă trăiri conștiente. Exemple de trăiri conștiente sunt trăirile vizuale (clasic este exemplul culorilor), auditive, tactile, olfactive, gustul, durerea, ceea ce se numește *imaginația mintală* (producerea în minte de imagini care nu au corespondent în realitate, dar a căror percepție o putem trăi), gândirea conștientă, simțul eului.<sup>265</sup> Toate acestea constituie “cealaltă față” a minții noastre, prima fiind cea psihologică.

Absența unei explicații reductive a fetei fenomenologice a minții noastre a fost denumită *lacună explicativă*.<sup>266</sup> Terminologia aceasta ar sugera un soi de lipsă care ar putea fi cumva eliminată. Nu este nimic de eliminat însă aici dacă

<sup>265</sup> Chalmers, *The Conscious Mind*, pp. 6-10.

<sup>266</sup> De către J. Levine în 1983 (cf. Chalmers, *The Conscious Mind*, p. 107).

acceptăm ideea imposibilității supervenienței logice. Faptele ce țin de conștiința fenomenologică sunt fapte suplimentare în raport cu faptele de ordin fizic.

Din acest punct, la prima vedere, te-ai aștepta ca discuția să ia în filosofie turnura pe care a mai luat-o în trecut. Faptele ce țin de conștiința fenomenologică ar ține de o zonă inaccesibilă investigației științifice. "Misterele" conștiinței ar fi cumva prin natura lor inaccesibile. Ar fi vorba despre mistere nu în sens metaforic, ci chiar la propriu. David Chalmers crede însă că este posibil un alt drum.

Cuvântul peste care nu trebuie trecut prea repede în toate argumentele discutate mai sus este "logic". Ființele zombi sunt logic posibile. Asta nu înseamnă că lucrurile nu se schimbă dacă luăm în considerare nu lumile alternative logic posibile, ci doar alternativele naturale. Punctul de vedere susținut de Chalmers este că există superveniență naturală a conștiinței fenomenologice, dar nu o superveniență logică a acesteia peste mulțimea faptelor fizice. Chalmers numește această viziune filosofică *dualism naturalist*.<sup>267</sup>

Implicația poziției de mai sus este însă o *expandare* a naturii. Faptele ce țin de conștiința fenomenală vin să se adauge celor fizice. Cu acest concept lărgit de natură se poate, mai departe, opera pentru a produce explicații *naturale* ale conștiinței.

Dacă o luăm pe acest drum, ne așteaptă la capăt o surpriză de proporții. Argumentele de mai sus au o structură izbitor de asemănătoare cu acelea ale criticilor inteligenței artificiale, mai precis cu cele ale lui Searle. Plasate însă în carcasa conceptuală creată de superveniență, ele conduc la ideea că sunt totuși posibile explicații naturale ale conștiinței fenomenale. Dacă aceste explicații sunt posibile, atunci este posibilă, cel puțin în principiu, realizarea de sisteme artificiale

<sup>267</sup> Chalmers, *The Conscious Mind*, p. 123 ș.u.

inteligente într-un sens foarte tare și anume sisteme care să aibă și trăirile noastre.<sup>268</sup>

Care este cheia argumentului în favoarea susținerii de mai sus? Totul ar ține, crede Chalmers, de implementarea unui sistem computațional. Pentru a realiza un astfel de sistem pornim de la abstracția sa (vezi capitolul despre *noile unelte* pentru acest sens al conceptului de *abstracție*). Ca să ajungă să fie un sistem real, un sistem abstract trebuie însă *implementat*. Structura cauzală a sistemului real trebuie să reflecte structura abstractă. Abordarea computațională oferă însă suficiente resurse pentru a “captura” structura cauzală relevantă.

În cel mai rău caz, ar trebui recurs la experimente filosofice imaginare de genul celui în care neuronii dintr-un creier care are trăiri conștiente sunt treptat înlocuiți cu mici demoni care au însușirea de a duce la bun sfârșit aceleași sarcini ca și neuronii. Înlocuirile sunt făcute treptat. Nu există nici un motiv pentru a crede că trăirile dispar.<sup>269</sup>

Perspectiva deschisă mai sus este cea unui funcționalism nonreductiv.<sup>270</sup> Organizarea funcțională a sistemului computațional determină trăirile conștiente. Conștiința nu se reduce însă logic la organizare funcțională.

Din perspectiva cărții de față acest argument este important deoarece coroborează poziția conform căreia sistemele computaționale nu sunt o simplă instanțiere a unor sisteme logice. În aceste condiții, ar fi și surprinzător ca numai metodele

---

<sup>268</sup> Chalmers, *The Conscious Mind*, pp 314-315

<sup>269</sup> Pentru argumentarea pe care o implică acest experiment gândit vezi Chalmers, “Absent Qualia, Fading Qualia, Dancing Qualia”, cap.7 din *The Conscious Mind*, pp. 247-275.

<sup>270</sup> Chalmers, *The Conscious Mind*, p 274.

logice să permită realizarea de sisteme computaționale. De aici caracterul genuin al *noilor unelte*.

## În jungla bibliografiei

Fișele făcute de-a lungul anilor, cu notițe și extrase pregătite pentru a scrie acest text, au luat aspectul unui labirint care mă deprimă. Nici forma electronică adoptată mai nou nu pare a fi mai puțin întortocheată. De aceea, a le transforma într-o listă alfabetică nu mi se pare cu rost. Voi încerca să găsesc firul sau firele Ariadnei prin acest labirint, în așa fel încât cititoarea sau cititorul să poată alege calea cea mai potrivită.

Mai precis, bibliografia de mai jos este construită pe ideea "capetelor de fir": un număr relativ mic de lucrări citate, dar care o dată consultate permit descoperirea altor lucrări (citate de către autorii lucrărilor menționate aici) și așa mai departe. Sunt menționate și unele "rădăcini" (*rădăcini* în sensul că nu mai conduc către altceva prin această metodă a *arborelui bibliografic*; Ryle *The Concept of Mind* este exemplul limită de "fundătură": nu are note de subsol și, în text, trimite practic doar la Descartes). Nu sfătuiesc pe nimeni să ia aceste *rădăcini* drept puncte de plecare. Sunt nenumărate fire care pleacă din ele, se desfac și se încurcă în așa fel încât nu se știe unde ajungi prin junglă și dacă mai dai peste dezbaterele vii în prezent. Ca să te descurci ar trebui să *refaci istoria*<sup>271</sup> acestor căutări, lucru care este în afara obiectivelor pe care ni le-am propus aici.

În altă ordine de idei, un defect *local* al listelor bibliografice este lipsa, în bibliotecile din țară, a titlurilor la care se face referire. Există două remedii ale acestor lacune. Primul este recursul la fonduri de carte relativ mici, dar selecționate cu grijă. Nu veți găsi toate titlurile, dar veți găsi aproape toate ideile esențiale. De exemplu, cei care au acces la Biblioteca Facultății

---

<sup>271</sup> Cred că *echilibrul prin răsfângere*, cum se spune uneori, este cheia unei bune istorii: să reiei drumul de la rădăcini, dar cu o bună cunoaștere a "capetelor de fir" (a rezultatelor din prezent), să reconstitui rădăcinile și apoi să faci mișcări de balans, pentru a echilibra înțelegerea, prin prezent, a trecutului cu iluminarea, prin trecut, a prezentului. Investigația care mi se pare cea mai inspirată din acest punct de vedere este cea a lui Jan Łukasiewicz din *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic* (1957).

de Filosofie a Universității din București pot consulta un fond de cărți de filosofie minții și știința cogniției al cărui nucleu a fost constituit, începând cu 1990, prin donațiile profesorului Radu J. Bogdan de la Tulane University (SUA).<sup>272</sup>

Pe de altă parte, entuziasmul organizatorilor Cercului de știința cogniției din Facultatea de Filosofie, Gabriel și Mihai Vacariu, a contribuit considerabil la depășirea barierelor bibliografice. Multe dintre titlurile menționate mai jos au fost găsite și reunite de către ei în biblioteca cercului.

Din păcate, nu toate titlurile de mai jos sunt disponibile în cadrul Universității din București. Multe lucrări se găsesc fie în biblioteca CEU la Budapesta, fie în biblioteci din Paris, Oxford sau Tulsa. Ar fi și greu să-i numesc aici pe toți cei care m-au ajutat să ajung la ele, dar le sunt profund îndatorat.

Am adăugat mici notițe explicative. Acestea pot da o idee cu privire la conținut, dar și cu privire la ce anume ar putea înlocui foarte bine titlul respectiv. De exemplu, o carte despre limbajul Pascal poate fi substituită cu alta. Cazul cărților de filosofie, este drept, e mai dificil, dar trebuie observat că oricum trebuie croit un drum prin jungla bibliografiei. Nu se poate citi tot.

Al doilea remediu la insuficiențele de ordin bibliografic este recursul la Internet. Filosofia minții și știința cogniției sunt excelent reprezentate pe Internet. Aș zice chiar că există un contrast, din acest punct de vedere, cu alte arii ale filosofiei (cu excepția textelor clasice). Cea mai bună ilustrare a acestei idei o constituie, cred, posibilitatea de a urmări, la zi, discuțiile care preced publicarea unui articol în *Behavioral and Brain Sciences*. Mari dezbateri din filosofia minții și știința cogniției s-au desfășurat prin intermediul acestei reviste (vezi

---

<sup>272</sup> Dacă veți deschide tratatul de știința cogniției editat de Michael I. Posner (*Foundations of Cognitive Science*, Cambridge[S.U.A.], MIT Press, 1989, 888p., cota III 1008) veți da peste următoarele rânduri, ce trădează entuziasmul generat de anul 1989: "Biblioteca Facultății de Filosofie, și prin ea, prietenilor întru cogniție. Un omagiu și o modestă contribuție noilor timpuri politice și intelectuale. Radu J. Bogdan, București, iunie 1990." Fiecare carte pe care profesorul Bogdan a adus-o Bibliotecii Facultății de Filosofie, an de an începând cu 1990, stă mărturie pentru generozitatea sa și stabilitatea amintitei prietenii întru cogniție.



<http://www.cogsci.soton.ac.uk/bbs/Archive> pentru arhivele revistei; tot la această adresă veți găsi și cuprinsul revistei, precum și articole propuse spre publicare și care, după regula celebră a revistei, sunt comentate în mod deschis de către membrii comunității de specialitate).

Internetul, în măsura în care este accesibil, face ca decalajul informațional dintre Est și Vest să fie în mare măsură depășit. Dacă este completat și cu un mic laborator care să permită desfășurarea investigației în direcția *filosofiei computaționale* situația tinde să se apropie de normal în zona această zonă.

Volumul de față presupune o anumită familiarizare cu logica. Referirile din text sunt indirecte și netehnice. Câteva trimiteri la o literatură mai tehnică sunt incluse în continuare (v. Miroiu, *Logica filosofică*, p. 200 pentru o instructivă trecere în revistă a nivelului de tehnicitate al literaturii relevante).

La urmă, dar nu în cele din urmă, ar trebui adăugat că foarte mult soft pentru știința cogniției poate fi găsit pe Internet (vezi pagina de web a Facultății de Filosofie a Universității din București la <http://www.fil.unibuc.ro> pentru legăturile cu centrele de unde se pot obține asemenea programe).

### **Bibliografii, ghiduri, dicționare și enciclopedii**

- **CHALMERS, David.** *Contemporary Philosophy of Mind: An Annotated Bibliography.* Pe Internet la <http://www.u.arizona.edu/~chalmers/biblio.html> [Este probabil cea mai bună bibliografie de filosofia minții. Conține sute de titluri. Vechea adresă pe Internet era la [ling.ucsc.edu/~chalmers](http://ling.ucsc.edu/~chalmers). Schimbarea de adrese pe Internet este, desigur, o problemă. La nevoie, se poate căuta însă adresa dorită folosind cuvinte-cheie.]
- **CRAIG, Edward (editor general) și Luciano Floridi (editor consultant).** *Routledge Encyclopedia of Philosophy.* Londra, Routledge, 1998. [Disponibilă și pe CD-ROM. Cea mai vastă enciclopedie de filosofie conține ample articole de filosofia minții, precum și bibliografii. Este foarte ușor de căutat prin ea în formatul electronic și oferă o

posibilitatea de a arunca o privire asupra contextului mai vast, inclusiv istoric, al dezbaterilor din filosofia minții.]

- **DENNETT, Daniel.** *Publications List* [URL = <http://ase.tufts.edu/cogstud/pubpage.htm> ]. [Ghidul publicațiilor lui Dennett, unul dintre autorii cei mai importanți din filosofia minții și știința cogniției. Numeroase texte sunt publicate electronic.]
- *Dictionary of Philosophy of Mind* [URL = <http://www.artsci.wustl.edu/~philos/MindDict/main.html>] [Explicații ale termenilor fundamentali în filosofia minții și biografii. Puteți găsi și biografia lui Brentano, dar și pe cea a lui Rodney Brooks (unul dintre creatorii lui Cog) sau pe cea a lui Jerry Fodor.]
- **FIESER, James (editor) și Bradley Dowden (editor asistent).** *Internet Encyclopedia of Philosophy*. URL = <http://www.utm.edu/research/iep/> . [Conține și o colecție de texte. Articolele sunt relativ scurte. Menționează reperele cele mai importante. De pildă, articolul despre *behaviorism* (din 1997) trasează distincția dintre *behaviorism hard* și *soft* (inclusiv *metodologic*), legătura dintre *behaviorism* și *verificaționismul pozitivist*, *behaviorismul logic* al lui Ryle, precum și criticile posibile la adresa *behaviorismului*.]
- **GUTTENPLAN, Samuel (ed.).** *A Companion to the Philosophy of Mind*. Oxford, Blackwell, 1994. [Articolele sunt ordonate alfabetic și sunt însoțite de bibliografii substanțiale. Ghidul acesta de filosofia minții include autoprezentări ale unor mari personalități ale domeniului: Chomsky, Davidson, Dennett, Dretske, Fodor, Putnam, Searle, Stalnaker. Există și articole consacrate lui Quine, Ryle, Turing și Wittgenstein. Guttenplan a scris o introducere amplă pentru cei nefamiliarizați cu filosofia minții.]
- **HODGES, Andrew.** <http://www.turing.org.uk/turing/Turing.html> [Este locul de pe Internet de unde se poate porni explorarea științei cogniției. În fond, inteligența artificială începe cu Turing. Ca ghid, acest site de pe Internet este prețios prin orientarea pe care o oferă în opera lui Turing.]

- **HOUDÉ, Olivier și Daniel Kayser, Olivier Koenig, Joëlle Proust, François Rastier (coordonatori). *Vocabulaire de sciences cognitives*. Presses Universitaires de France, 1998.** [În introducere sunt prezentate pe scurt disciplinele acoperite de către dicționar: neuroștiința, psihologia, inteligența artificială, lingvistica și filosofia. Este schițat și drumul parcurs de către știința cogniției în Franța. Articolele sunt de tip enciclopedic, în sensul că tratarea subiectului este făcută din perspectiva mai multor discipline. Bibliografiile sunt extrem de prețioase pentru că permit urmărirea literaturii relevante și în limba franceză.]
- **LAU, Joe. *The International Directory of On-Line Philosophy Papers*. URL = <http://hkusuc.hku.hk/philodep/directory/>.** [Aceasta este o bază de date cu legături către diversele locuri de pe Internet unde se găsesc texte de filosofie. Avantajul acestui ghid este că lucrările sunt clasificate și nu trebuie căutat prin toată lumea pentru a le găsi. Evident, este mai mult decât o bibliografie tradițională, deoarece poți să ajungi rapid și la texte.]
- **LYCAN, William G. "Philosophy of Mind" în *The Blackwell Companion to Philosophy*. Oxford, Blackwell, 1996, pp. 167-197.** [Cuprinde următoarele secțiuni: raportul corp-minte, behaviorism, teoria identității, funcționalismul psihologia cognitivă, inteligența artificială și modelul computațional al minții, funcționalism homuncular și alte teorii, qualia și conștiința, intenționalitatea, ipoteza limbajului gândirii, instrumentalismul, eliminativism și neurofilosofie, problemele psihologiei populare.]
- **MALATESTI, Luca. *Philosophy of Mind*. URL = <http://lgxserver.uniba.it/lei/MINDE/Homemine.htm>.** [Ghid al resurselor de filosofia minții de pe Internet. Are baze de date cu legături către bibliografii, pagini de web, anunțuri ale manifestărilor științifice, reviste, lucrări de referință, instituții, texte. Deosebit de utile mi s-au părut informațiile cu privire la cărțile recent apărute și cele care urmează să apară. Sunt indicate inclusiv legăturile către cărțile aflate în lucru. De exemplu, cunoscutul filosof John Perry elaborează o carte intitulată *Knowledge, Possibility and Consciousness* (pe baza prelegerilor Nicod) și o variantă preliminară poate

di consultată pe web la URL = [www-csli.stanford.edu/~john/NICOD/VERSIONS/0-book/node1.html](http://www-csli.stanford.edu/~john/NICOD/VERSIONS/0-book/node1.html) Prezența variantelor de lucru pe Internet spune mult și despre diferențe de mentalitate. Autorii nu se ascund într-o peșteră, de unde ies cu cartea gata și așteaptă laudele. Cartea este citită pe parcurs de colegi, se adună observații; autorul încearcă să respingă obiecțiile; dacă nu poate să o facă, atunci modifică textul său. În rezumat, toate aceste resurse și ghidurile către ele țin de standarde mai degrabă neobișnuite pe la noi.]

- **NANI, Marco și Massimo Marraffa (editori). *A Field Guide to the Philosophy of Mind*. URL = <http://www.uniroma3.it/kant/field/index.html> .** [În 1999 ghidul acesta este încă în construcție. Cuprinde eseuri destul de întinse pe marile probleme ale filosofiei minții plus bibliografii adnotate. Dintre temele deja acoperite menționăm: filosofia kantiană a minții, argumentul cunoașterii (experimentul lui Jackson etc.), internalism/externalism, limbaj și gândire, limbajul gândirii, metafizica minții, Pământurile îngemănate (vestitul experiment gândit propus de Putnam). Multe eseuri sunt semnate de autori din țările de limbă engleză. Există însă și eseuri semnate de filosofi din Italia. De exemplu, eseu despre contrastul dintre internalism și externalism este semnat de Alberto Voltolini (Universitatea din Torino – Vercelli). Exemplu de externalism este poziția lui Putnam (din 1975 încoace). Internaliști sunt autori ca Fodor, Chomsky, Jackendoff.]
- **Philosophie des Geistes [URL = <http://www.philo.de/Philosophie-Seiten/themen/geist.htm> ]** [Ghid de legături pe Internet, inclusiv pentru Germania. Trimiteri utile la cărțile de filosofia minții și știința cogniției apărute în Germania.]
- **RAGGETT, Jenny și William Bains. *Artificial Intelligence from A to Z*. Londra, Chapman and Hall, 1992.** [Un dicționar elementar al inteligenței artificiale. Există și o traducere în maghiară (Budapesta, 1994). Număr foarte mare de articole care explică termeni filosofici – un semn al relației aparte dintre filosofie și inteligența artificială.]

- **WILSON, Robert A. și Frank Keil (editori).** *The MIT Encyclopedia of Cognitive Science*. The MIT Press, 1999. [Cât timp a fost în construcție, înainte de a apărea în volum și pe CD, enciclopedia a fost accesibilă liber pe Internet, până în iunie 1999. Aceasta face parte desigur din efortul autorilor de strânge cât mai multe reacții critice față de proiect și de texte ca atare. Din varianta finală sunt accesibile liber (URL = <http://mitpress.mit.edu/MITECS/>) rezumatele articolelor. Ele sunt practic ca un mic dicționar de știința cogniției. Marele lor avantaj este însă altul. Rezumatele sunt însoțite de bibliografii substanțiale.]
- **ZALTA, Edward N. (editor).** *Stanford Encyclopedia of Philosophy*. URL = <http://plato.stanford.edu/> . [Articolele de filosofia minții (câte sunt scrise până în acest moment) nu sunt foarte numeroase, dar sunt bine structurate și utile. De exemplu, articolul despre știința cogniției este semnat de către Paul Thagard. Începe cu un scurt istoric, în care George Miller, John McCarthy, Marvin Minsky, Allen Newell, Herbert Simon și Noam Chomsky sunt prezentați ca fondatori ai domeniului. Continuă cu o secțiune despre reprezentări și computații (ipoteza centrală a științei cogniției este că gândirea poate fi înțeleasă în termenii structurilor reprezentationale și ai procedurilor computaționale care le prelucrează – paralelismul este cu structurile de date și algoritmi din programare). A treia secțiune este dedicată filosofiei. Accentul cade pe posibilitatea de a da din perspectivă computațională soluții unor vechi probleme filosofice. Sunt sistematizate și criticile la adresa științei cogniției. O bibliografie și o listă de legături pe Internet încheie articolul.]

### Antologii și introduceri

- **ANDLER, Daniel.(ed.)** *Introduction aux sciences cognitives*. Paris, Gallimard, 1992. [Cuprinde o introducere în problemele calculului și ale reprezentării, creierului, modularității minții, limbajului și comunicării, formării conceptelor și teoriilor, sistemelor sociale văzute din

perspectivă cognitivă. Este rodul cooperării unor cercetători din Franța<sup>273</sup>, Canada, Anglia, Australia și SUA. ]

- **Block, Ned (editor). *Readings in Philosophy of Psychology*, Vol. II. Harvard University Press, 1981.** [Cuprinde texte importante, scrise de Chomsky, Fodor, Katz ș.a., precum și dezbateri în probleme celebre, ca - de pildă - cea a ideilor înnăscute.]
- **BOGDAN, Radu J. (editor). *Mind and Common Sense*. Cambridge University Press, 1991.** [Este un volum colectiv, care reunește studii despre psihologia simțului comun, semnate - între alții - de către Paul M. Churchland, R. Cummins, Colin McGinn, Alexander Rosenberg, Jonathan Bennett.]
- **BRADDON-MITCHELL, David și Frank Jackson. *The Philosophy of Mind and Cognition*. Oxford, Blackwell, 1996.** [Introducerea pornește de la discuția despre dualism și behaviorism; trece apoi la teoria identității și funcționalism, conținut mental și explicarea comportamentelor. Cuprinde părți foarte utile despre experimentele imaginare propuse de către Frank Jackson. Autorii combină o introducere pentru cei care explorează pentru prima dată filosofia minții cu dezvoltarea și apărarea propriei viziuni în filosofia minții. S-ar putea să greșesc, dar impresia mea a fost că volumul trădează un anumit izolaționism filosofic - legăturile cu alte discipline interesate de cogniție sunt practic tăiate. Autorii sunt centrați pe funcționalismul lor care privilegiază ideea că rolurile pe care simțul comun le asociază stărilor mentale conferă înțeles termenilor care se referă la stările mentale.]

---

<sup>273</sup> În Franța, din perspectiva filosofiei, centrul unde se cultivă investigațiile de filosofia minții și știința cogniției este CREA (Centre de Recherche en Epistémologie Appliquée). CREA este totodată vârful de lance al filosofiei analitice în Franța. Directorul CREA este Daniel Andler, profesor la Paris X Nanterre. Dintre membrii Centrului, doamna Proust a fost editoarea părții de filosofie a *Vocabulaire de Sciences Cognitives*. Pascal Engel, profesor la Sorbona, este cercetător în cadrul Centrului. Membri ai Centrului sunt și Pierre Jacob, François Récanati, E. Pacherie. Centrul a publicat o serie întreagă de caiete, inclusiv unul dedicat filosofiei minții (vezi URL = <http://www.polytechnique.fr/laboratoires/crea/F/cahiers.html>).

- **CARALP, Evelyne și Alain Gallo. *Sciences cognitives et psychanalyse*. Toulouse, Milan, 1998.** [Introducere concisă, dar foarte utilă. Cuprinde elemente de psihologie cognitivă, neuroștiință, inteligență artificială și psihanaliză.]
- **CHURCHLAND, Paul M. *Matter and Consciousness*. Ediție revăzută. The MIT Press, 1996.** [Ediția originală a apărut în 1984. Începe cu analiza a patru probleme: ontologică, semantică, epistemologică și metodologică. Prezintă rezultatele inteligenței artificiale și ale neuroștiinței. Expunere foarte clară și concisă. Introducerea în neuroștiință este prețioasă pentru filosofi. Este probabil modelul de introducere în filosofia minții din perspectiva științei cognitive.]
- **DENNETT, Daniel. *Kinds of Minds*. Londra, Orion, 1997.** [Prima ediție 1996. Este tradusă în limba română la Editura Humanitas. O ilustrare excelentă a importanței întrebărilor pentru filosofie. Foarte utilă pentru cei care vor să înțeleagă ce anume înseamnă *intenționalitate* în filosofie: a fi despre, a fi orientat către altceva, a avea un obiect. Reprezentările au caracter intențional: sunt orientate către un referent.]
- **ENGEL, Pascal. *Introduction à la philosophie de l'esprit*. Éditions La Découverte, 1994.** [Introducere în universul cauzelor mentale, creștințelor, interpretării umane și animale, conținutului și imaginilor mentale, identității personale și al conștiinței. În Franța există și o altă filosofie decât cea prezentă în jumale. Această filosofie, de factură analitică, are printre reprezentanții săi de marcă pe Pascal Engel.]
- **FLONTA, Mircea. *Cognitio. O introducere critică în problema cunoașterii*. București, ALL, 1994.** [Deși teoria cunoașterii este mai puțin implicată direct în știința cognitive decât filosofia minții, zonele de suprapunere sunt totuși semnificative (atitudini propoziționale, discuția despre empirism *versus* raționalism în învățarea limbii și, fără doar și poate, însuși conceptul de cunoaștere). Mircea Flonta expune cu obișnuita sa limezime și sistematicitate aceste subiecte. A se vedea, în special, "Înnăscut și dobândit" (pp. 105-112), inclusiv pentru terminologie (*minte, nativism, știința cognitive*). Am deviat de la terminologia din *Cognitio*

în cazul lui *belief*, pe care Mircea Flonta îl redă prin *opinie* (aici *crezământ*), și *proposition*, pe care Mircea Flonta îl traduce prin *judecată* (aici *conținut propozițional*). Pentru o prezentare enciclopedică a teoriei cunoașterii a se vedea Jonathan Dancy și Ernest Sosa (editori) *A Companion to Epistemology* (Oxford: Blackwell, 1992), din care un prim volum a fost tradus în limba română la Editura Trei de către Dan Bărbulescu, Corina Ștefanov, Gheorghe Ștefanov și Anabela Zagura. Interesant este că Dancy și Sosa au un articol despre *alte minți*, în vreme de Guttenplan *Companion to the Philosophy of Mind* nu conține un articol separat pe această problemă tradițională în filosofia minții.]

- **GARDNER, Howard. *The Mind's New Science. A History of the Cognitive Revolution*. 1985 [Cu o postfață a autorului la ediția din 1987]** [O sursă de date foarte prețioasă, mai ales atunci când este vorba despre istoria științei cogniției. Probabil că și astăzi este lectura cu care ar trebui începută familiarizarea cu știința cogniției.]
- **GLOVER, Jonathan (editor). *The Philosophy of Mind*. Oxford University Press, 1976.** [Problematika antologiei este divizată în patru arii: descriere și clasificare, modele ale minții, problema corp-minte, problema identității personale. Glover crede că studiile lui Derek Parfit ("Personal identity", pp. 142-162) și Bernard Williams ("The self and the future", pp. 126-141) reprezintă pentru oricine adevărate teste ale interesului pentru filosofie: dacă ești în stare să le comentezi, atunci ești apt/aptă pentru filosofie. Oricum ar sta lucrurile, aceste investigații sunt, ce-i drept, veritabile modele de utilizare a experimentelor imaginare în filosofie. Ele fac apel însă la experimente propuse Shoemaker, așa că nu este rău să vă uitați și prin scrierile acestui filosof.]
- **GOLDMAN, Alvin I.(ed.) *Readings in Philosophy and Cognitive Science*. The MIT Press, 1993.** [Acoperă tematica impactului științei cogniției asupra filosofiei în general. Capitolul al treilea este dedicat filosofiei minții. Ultimul capitol, cel consacrat fundamentelor conceptuale, cuprinde texte foarte importante pentru bazele științei cogniției, inclusiv faimoasa dispută dintre Fodor și Pylyshyn, pe de o parte, și Smolensky de cealaltă parte. Simplul fapt



că volumul are peste 850 de pagini spune ceva despre explozia interesului pentru știința cogniției în filosofie.]

- **HANSEN, Frank-Peter (editor). *Philosophie von Platon bis Nietzsche*. Berlin, Directmedia, 1998.** [Această antologie pe CD-ROM, tipărită, s-ar întinde pe zeci de volume. Are însă și lipsuri, poate tocmai din perspectiva filosofiei minții. De pildă, din Aristotel sunt incluse doar *Organon*-ul, *Fizica*, *Metafizica* și *Etica nicomahică*. Avantajul este însă dat de posibilitatea de a căuta electronic. De exemplu, dacă vom căuta cuvântul *Physiognomie* (și termenii înrudiți), îl vom găsi nu doar la Hegel, ci și la Aristotel, Paracelsus, Mendelssohn, La Mettrie, Rousseau, Herder, Lessing, Jean Paul, Lange, Hartmann, Schopenhauer și alți autori. Evident, referirile trebuie selectate. Este vorba uneori de "fizionomie" în sensul uzual al termenului. De asemenea, nu înseamnă că discuția are aceeași greutate ca la Hegel ș.a.m.d. Avantajele textelor electronice sunt însă considerabile, chiar și atunci când este vorba despre căutări foarte simple.]
- **HAUGELAND, John (editor). *Mind Design*. The MIT Press, 1981.** [Una dintre antologiile care au contribuit la răspândirea științei cogniției. Cuprinde, între altele, studiile lui Newell și Simon despre simboluri și căutare, Minsky despre reprezentarea cunoașterii, Marr despre inteligența artificială, Fodor despre solipsismul metodologic, Davidson despre mintea materială, precum și reflecțiile critice ale lui Dreyfus și Searle.]
- **HEIL, John și Alfred Mele (editori). *Mental Causation*. Oxford, Clarendon Press, 1995.** [Davidson precizează, în "Thinking Causes" (pp. 3-17) modul în care a introdus ideea de *superveniență* în filosofia minții și raporturile cu felul în care ea a fost folosită în filosofia morală de către Moore și Hare. Pentru Davidson un predicat *p* este supervenient peste o mulțime de predicate *S* dacă *p* nu poate servi drept trăsătură distinctivă a unei entități fără ca să poată fi folosit în același scop un predicat din *S*. Cazurile interesante sunt cele în care *p* nu poate fi reduc la predicatele din *S*. Davidson precizează apoi atitudinea sa față de idei ale lui Kim, Fodor și alții. Tot în acest volum Frank Jackson și Philipp Petit semnează "Some content is narrow"

(pp. 259-282). H. W. Noonan scrie despre gânduri dependente de obiecte (despre disputa internalism-externalism, cu alte cuvinte).]

- **HEIL, John. *The Philosophy of Mind*. Routledge, 1998.** [Este o introducere din perspectivă *metafizică* în filosofia minții.]
- **HONDERICH, Ted. (editor). *Philosophy Through Its Past*. Pelican Books, 1984.** [Filosofia actuală a minții și, cu atât mai mult, știința cogniției au memoria scurtă. Scrierile din jurul anilor 1975-1980 sunt deja *clasice*, ca să nu mai vorbim de textele dintre 1950-1975, care în mod sigur sunt clasice. Asta nu înseamnă că nu este interesant să ne aruncăm o privire mai atentă în trecutul mai îndepărtat. Reputata colecție de eseuri editată de Honderich selectează nouăsprezece nume din istoria filosofiei (de la Platon la Sartre). O treime dintre eseuri au o tematică centrată pe probleme de filosofia minții. Anthony Kenny îl studiază pe Toma d'Aquino din perspectiva conceptului de *intenționalitate*. Pune în evidență rosturile conceptului aristotelic de *formă* pentru înțelegerea intenționalității (formele pot fi în natură, dar pot fi și în minte). Gândim în forme universale, dar în *contexte* senzoriale. Bernard Williams scrie despre *cogito* și *sum* la Descartes. Ceea ce mă derutează oarecum în textul lui Williams este felul în care el se centrează pe structura inferențială, nu pe experimentele imaginare când este vorba de domeniul evenimentelor de gândire (*thought-events*). J. L. Mackie discută teoria percepției la Locke. Din cunoscuta carte a lui Barry Stroud despre Hume este ales capitolul despre acțiune, rațiune și pasiune. Alasdair MacIntyre încearcă să găsească la Hegel un sprijin în favoarea ideii că fizica nu poate oferi o explicație cauzală a reflecției raționale și a conștiinței. Hegel se raporta polemic la fiziognomie și frenologie. Obiecțiile hegeliene, dacă ar sta în picioare, s-ar aplica și neuroștiinței din ziua de azi. Dacă "vedem" că o persoană are o expresie facială tristă, lucrul acesta se explică hegelian prin convenții culturale. În altă cultură, ar trebui să "vedem" conform convențiilor locale. Nu este nimic de corelat cu ceva intern. Ayer atrage atenția asupra analizei minții la Russell (fără să focalizeze, este adevărat,

întreaga discuție pe problemele minții). Russell credea că sunt două moduri de a ne referi la un obiect: putem să-l *indicăm* (în engleză: *point*) sau să-l descriem. Pentru ca *indicarea* să funcționeze este nevoie ca obiectul să existe. Iar existența sa este stabilită prin datele senzoriale din prezent. Putem oare indica și un obiect de care doar ne amintim? Până la urmă, Russell avea să ajungă la concluzia că la obiectele din memorie avem acces tot indirect. (Vezi și Russell *The Analysis of Mind* mai jos, în secțiunea despre problemele tradiționale.) În sfârșit, Danto pune în evidență felul în care, dincolo de expresiile strălucitoare ale lui Sartre, de genul *L'enfer c'est les autres*, se găsește o rafinată sensibilitate filosofică pentru problema altor minți. Starea de *rușine*, de pildă, nu este inteligibilă filosofic, din perspectiva lui Sartre, decât dacă este presupusă existența altor minți. Așadar filosofia minții este o prezență considerabil mai veche în istoria filosofiei decât ar sugera-o notele de subsol ale majorității scrierilor de specialitate din ziua de azi. Diferența cred că este dată de impactul științei cognitive. De aici și limitele tradiției, mai ales că o traducere a vechilor idei în cadru cognitivist, așa cum se face în antologia lui Honderich în cadru analitic, lipsește.]

- **HOOK, Sidney (editor). *Language and Philosophy*. New York, 1969. [Traducere în italiană (Roma: Armando, 1971)]** [Conține lucrările unui simpozion de la Universitatea din New York din 1968. Secțiunea dedicată raporturilor dintre lingvistică și filosofie se deschide cu intervenția lui Noam Chomsky. Chiar de la bun început Chomsky afirmă că lingvistica și filosofia sunt de nedespărțit. Dacă vom corela această afirmație cu cea a lui McCarthy despre legătura dintre inteligența artificială și filosofie și cu insistența aceluiași Chomsky în direcția unei legături între lingvistică și psihologia cognitivă, regăsim sistemul de raporturi care stă la baza științei cognitive ca demers multidisciplinar.]
- **JOHNSON-LAIRD, Philip. *The Computer and the Mind. An Introduction to Cognitive Science*. Ediția a doua. Fontana Press, 1993.** [O excelentă introducere în știința cognitive făcută din perspectiva psihologului. Cuprinde și o prezentare a teoriei modelelor mentale dezvoltate de

Johnson-Laird. Dacă vreau să prind ideea că "dulapul se află la dreapta frigiderului", atunci un item din model va fi pus la dreapta altui item. Ambiția este de a pune în evidență felul în care situațiile pot fi reprezentate în minte.]

- **LEVY, Bruno (redactor-șef) și Emil Servan-Schreiber (realizator). *Les Secrets de l'intelligence*. 2 CD-ROM. Paris, Hypermind [URL = <http://www.hypermind.com>] în colaborare cu Ubi Soft, 1997.** [Se vorbește cu groază în unele cercuri despre dispariția cărților. Cel puțin în cazul lucrărilor de popularizare a științei câștigul pare însă evident. Cele două CD-uri realizate la Paris prezintă știința cogniției cu o forță pe care nu ar putea-o avea nici o carte. Textul este combinat cu expuneri filmate. Poți face chiar experimente. Poți vedea faimosul robot Cog în acțiune. Îl poți asculta pe Rodney Brooks vorbind despre Cog. Grafica arată, de asemenea, fascinant și, desigur, creierul nostru – cum spun realizatorii – ne surprinde. Despre inteligența artificială și programele expert vorbește Herbert Simon. Imagistica mintală este prezentată de către Stephen Kosslyn. Iar conținutul, deși este cel al lucrărilor de popularizare, dat fiind caracterul multidisciplinar al științei cogniției, poate fi util studenților și chiar celor care se ocupă cu știința cogniției, dar nu s-au specializat în zona pe care o tratează secțiunea respectivă.]
- **McGINN, Colin. *The Character of the Mind*. Oxford University Press, 1997.** [Susține că distincția dintre filosofia psihologiei și filosofia minții este similară cu cea dintre filosofia lingvisticii și filosofia limbajului. Bibliografia folosită nu este întinsă, dar oferă un model de selecție exemplară: Armstrong, Davidson, Kripke, Nagel, Putnam, Chalmers, Burge, Fodor, Stich, Penrose, Searle, Parfit, Shoemaker, Chomsky și câteva alte nume cu o greutate similară.]
- **McGINN, Colin. *Problems in Philosophy. The Limits of Inquiry*. Oxford, Blackwell, 1993.** [Conține capitole dedicate conștiinței, eului, identității personale, dar și înțelesului, liberului arbitru și ordinii cauzale, cunoașterii și adevărului. Vezi și McGinn, *The Problem of Consciousness*.]

- **PERSEUS PROJECT.** [Pe Internet există Perseus Project. Cuprinde, între altele, o colecție de scrieri filosofice grecești: Proiectul Perseus are și textele în original, și traduceri, și: analize gramaticale, și dicționarul Liddel-Scott. A se vedea, de asemenea, *Thesaurus Linguae Graecae: Greek Texts* (colecție completă a textelor grecești în format electronic). Cu aceste surse nu se întâmplă, ca în antologia lui Hansen, să nu găsești Aristotel *De Anima*. Singurul regret este că utilitatea pentru filosofia minții este limitată.]
- **PLÉH, Csaba (ed.). *Kognitív tudomány*. Budapest. Osiris, 1996.** [În mod sigur aceasta este o carte utilă și pentru mulți cititori și cititoare de limbă maghiară din România. Antologia a fost pregătită pentru Colegiul Invizibil din Budapesta. În introducere, Pléh este de părere că presuposițiile filosofice ale computaționalismului sunt realiste (regulile sunt stări mentale ale subiectului), iar ale conexionismului sunt nominaliste (regulile există doar în mintea cercetătorului).]
- **POSNER, Michael I. (editor). *Foundations of cognitive science*. The MIT Press, 1989.** [Trecere în revistă amplă a științei cogniției. Capitolul omonim – cel introductiv – este semnat de Herbert A. Simon și Craig A. Kaplan. Simon este unul dintre părinții inteligenței artificiale și ai științei cogniției: o personalitate al cărei prestigiu trece dincolo de granițele acestor domenii (Simon este laureat al premiului Nobel pentru economie).]
- **ROSENTHAL, David M.(ed.) *The Nature of Mind*. Oxford University Press, 1991.** [Antologia are cinci părți: (1) problemele minții; (2) eul și alte minți; (3) minte și corp; (4) natura minții (gândire, percepție, conștiință, eu și persoană); (5) explicația psihologică]
- **SMITH, Peter și O. R. Jones. *The Philosophy of Mind*. Cambridge University Press, 1987.** [Partea întâi examinează dualismul. Partea a doua este consacrată atitudinilor propoziționale și evaluării teoriei funcționaliste a minții. Partea a treia este consacrată senzațiilor, gândirii și libertății.]

## Probleme tradiționale în filosofia minții

Jerry Fodor scrie că atunci când era în primii ani de facultate filosofia minții avea două părți mari: problema raportului corp-minte și problema altor minți.<sup>274</sup> Experiența mea a fost similară.<sup>275</sup> Acestea și sunt probleme centrale în discuțiile tradiționale de filosofia minții. Nu sunt însă și singurele probleme discutate de-a lungul istoriei filosofiei minții.

- **DENNETT, Daniel. Recenzie la Karl R. Popper și John C. Eccles *The Self and Its Brain* in *The Journal of Philosophy* LXXVI, 2/1979, pp. 91-97.** [Dennett susține că volumul lui Popper și Eccles are două defecte majore: autorii nu sunt coautori – părțile scrise separat de cei doi sunt discordante, iar discuția dintre ei are loc înainte ca ei să le scrie și nu rezolvă diferențele; autorii ignoră literatura recentă. Dennett avea să folosească totuși, ulterior, sugestia lui Popper că noi punem ideile să moară în locul nostru – v. Dennett, *Kinds of Minds*, p.116.]
- **DESCARTES, René. *Méditations métaphysiques*. Presses Universitaires de France, 1956. [Ediția a doua: "Quadrige", 1988]**
- **FELDMAN, Fred. "Kripke and the identity theory" in *The Journal of Philosophy* LXXI, 18/1974, pp. 665-676.** [Comentat de William G. Lycan. Prezentat la Simpozionul despre Kripke și materialiști din 1974.]
- **KIM, Jaegwon. *Supervenience and Mind*. Cambridge University Press, 1993.** [Discuțiile despre tradiționala problemă corp-minte sunt mai rafinate în prezent, grație recursului la superveniență. Acest instrument analitic este folosit și în debaterile despre conștiință. Kim este unul dintre autorii de referință în materie de superveniență.]

<sup>274</sup> Jerry Fodor, Autoprezentare în Guttenplan (ed.) *Companion*, p. 292.

<sup>275</sup> Cu precizarea că am aflat despre problema corp-minte la cercul organizat de profesorul Mircea Flonta. Nu exista un curs de filosofia minții, în Facultatea de filosofie. Evident, exista un curs de psihologie. Cât despre problema altor minți, de ea am aflat tangențial, din pricina interesului pentru Wittgenstein, cultivat – în anii '80 – la un cerc al cărui mentor era tot Mircea Flonta.

**KRIPKE, Saul A.** "Naming and Necessity" în D. Davidson și G. Harman (editori), *Semantics of Natural Language*. Ediția a doua. Reidel, 1972, pp. 253-355. [A se vedea critica din final la adresa teoriei identității (pp. 334-342).]

**KRIPKE, Saul A.** *Wittgenstein on Rules and Private Language*. Oxford, Blackwell, 1982. [Conține și un post-scriptum despre Wittgenstein și problema altor minți.]

**MORICK, Harold (ed.).** *Wittgenstein and the Problem of Other Minds*. McGraw-Hill, 1967. [Cuprinde, între altele, și critica lui Chihara și Fodor – din 1965 – la adresa analizei făcute de către Wittgenstein termenilor mentali ai limbajului obișnuit.]

**O'CONNOR, John (ed.).** *Modern Materialism: Readings on Mind-Body Identity*. New York, 1969. [Înțelegerea aprofundată a disecării teoriei identității, de la Saul Kripke și Frank Jackson la David Chalmers, nu se poate lipsi de consultarea surselor și evaluărilor critice reunite în acest volum: texte de J.J.C. Smart, Norman Malcolm, Richard Rorty, Richard Brandt, Jaegwon Kim și alții.]

**POPPER, Karl R.** *Knowledge and the Body-Mind Problem*. Londra, Routledge, 1994. [Tradusă în românește de Florin Lobonț la Editura Trei (1997).]

**POPPER, Karl și John C. Eccles.** *The Self and Its Brain*. Springer, 1977 [ediția originală; ediție revizuită 1881]. Routledge and Kegan Paul, 1984. [Ca în toată filosofia sa târzie, Popper distinge între trei lumi: *lumea 1* – obiecte și stări fizice -, *lumea 2* – cea a stărilor mentale și *lumea 3* – conținuturile obiective ale gândirii. Cele trei lumi interacționează. Interacțiunea dintre lumea 1 și lumea 2 este mai bine înțeleasă dacă luăm în considerare rolul lumii 3 – gazda ideilor, ipotezelor oamenilor – care face posibilă selecția critică a ipotezelor. V. și recenzia lui Dennett la carte.]

**PUTNAM, Hilary.** "Mind and Body" în *Reason, Truth and History*. Cambridge University Press, 1981. [Putnam expune pe scurt punctul său de vedere mai vechi, zis

*funcționalist*, dar apoi dezvoltă rezervele sale critice legate de caracterul calitativ al senzațiilor.]

- **PUTNAM, Hilary. *Mind, Language and Reality. Philosophical Papers, Volume 2. Cambridge University Press, 1975.*** [Studiile 14-22 sunt cele în care Putnam dezvoltă poziția sa *funcționalistă* în filosofia minții (de care ulterior avea să se distanțeze). Ca interval de timp, aceste studii acoperă deceniul al șaptelea și începutul deceniului al optulea al secolului al XX-lea. Cel mai discutat astăzi dintre studiile din acest volum este însă "The meaning of 'meaning'" (pp. 215-271). Faimoasă este întrebarea, pusă în acest studiu, dacă "înțelesurile sunt în cap?" (p. 223). Putnam propune experimentul său imaginar cu cele două Pământuri. Primul este chiar planeta noastră. Al doilea este un Pământ geamăn, dar pe care apa are același aspect ca și pe Pământul nostru dar o compoziție chimică diferită. Înțelesul nu poate fi "în cap", deoarece un omolog al meu de pe Pământul geamăn ar avea aceleași stări mentale ca și mine când ar gusta apa, cu toate că înțelesul termenului "apă" este diferit. După Putnam, înțelesul are și o componentă socială, și una legată de materialul din care sunt făcute obiectele pe care le luăm drept paradigme când fixăm înțelesul (or, acest material nu ne poate fi pe deplin cunoscut). Vezi, de asemenea, "Language and philosophy" (studiul cu care începe volumul) pentru interesanta interpretare a ceea ce înseamnă pentru Wittgenstein a avea un *concept* și motivele pentru care a avea un concept nu înseamnă a avea o imagine.]
- **ROBINSON, Howard. *Perception. Routledge and Kegan Paul, 1994.*** [Cartea expune și apără teoria percepției din perspectiva datelor senzoriale (*sense-datum theory of perception*). Discută, între altele, argumentele împotriva limbajului privat și interpretarea dată de Kripke lui Wittgenstein. Lucrarea este un bun "capăt de fir" pentru literatura consacrată filosofiei percepției.]
- **RUSSELL, Bertrand. *The Analysis of Mind. [Traducere italiană *L'Analisi della mente* (Roma: Newton Compton, 1969)]*** [Nu mă gândeam să introduc cartea lui Russell în bibliografie. M-am răzgândit pentru că Russell este



interesant din pricina rolului jucat de ei în descoperirea atitudinilor propoziționale. Passmore scrie că, pentru Russell, propozițiile sunt atomice și moleculare (John Passmore *A Hundred Years of Philosophy* (Harmondsworth: Penguin, 1968) pp. 234-235). Dacă  $P$  și  $Q$  sunt propoziții atomice, atunci " $P$  și  $Q$ " este moleculară. Adevărul ei depinde de *constituenți*. Dar "Cred că  $x$  este  $y$ "? Fie că " $x$  este  $y$ " e adevărată, fie că este falsă, eu tot s-ar putea să am (sau să nu am) o *relație de crezare* față de ea. A apărut o nouă "specie în grădina zoologică"? După Passmore, Russell a fost fericit să poată recurge la behaviorism în aceste cazuri: dacă mă întâlnesc cu  $x$ , atunci *reacția* mea este așa și așa. Passmore scrie că Russell a dat manuscrisul cărții lui J. B. Watson, pentru observații preliminare. Oricum ar fi, ideea de atitudine propozițională transpare în *Analysis of Mind*, chiar dacă tehnic vorbind Russell o va introduce mai târziu. Cartea începe cu idee că evenimentele tipic mentale sunt *crezămintele* și *dorințele*. Se referă apoi la conceptul de intenționalitate, preluat de Brentano din scolastica medievală (vezi și despre Brentano mai jos, în secțiunea despre reprezentări), dar respinge punctul de vedere al lui Brentano. Apoi, după ce scrie despre James, ajunge la behaviorismul lui J. B. Watson. Viziunea mare filosofică a lui Russell, în acest punct al operei sale, este că, de fapt, contrastul dintre materie și spirit este unul superficial. Există un substrat neutru din care sunt compuse atât corpul, cât și mintea. Legile fizice s-ar aplica astfel uniform. Russell este vestit pentru meandrele gândirii sale filosofice, așa că nu m-aș aventura să indic tranșant în ce direcție merg, până la urmă, analizele sale, pe parcursul ulterior al operei sale filosofice. Fapt este că, în carte, Russell are o poziție nuanțată, de pildă când este vorba despre introspecție. Există obiecții contra ei, dar nu toate faptele observate prin introspecție ar trebui scoase în afara cercetării științifice. Secțiunile despre percepție, senzație, imagini par a sta sub semnul semnificației analizei pentru o filosofie empiristă. Din punctul de vedere al minții, poziția lui Russell este că toate fenomenele psihice sunt formate numai din senzații și imagini. După ce discută limbajul și gândirea, Russell ajunge și la analiza în detaliu a crezămintelor.]

- **Ryle, Gilbert. *The Concept of Mind*. Penguin Books, 1970.** [Publicată pentru prima oară în 1949 de către Hutchinson]
- **SHOEMAKER, Sydney. *The First-Person Perspective and Other Essays*. Cambridge University Press, 1996.** [Eseul despre *perspectiva persoanei întâi* discută implicațiile adoptării acestei perspective asupra experimentelor imaginare la care se recurge în filosofia minții. Atacul lui Kripke la adresa tezei identității este analizat de autor prin această prismă. Este discutat pe larg experimentul lui Searle: înlocuirea de neuroni naturali, bucată cu bucată, cu neuroni artificiali. Ideea că am putea trece fără probleme, prin analogie, de la persoana întâi la persoana a treia este denumită de autor "una dintre buruienile cele mai încăpățânate" din grădina filosofiei (p. 174). Perspectiva persoanei întâi are presupuziții filosofice care sunt, de fapt, răspunsuri la probleme de filosofia minții, răspunsuri care nu pot fi testate din această perspectivă.]
- **SOLCAN, Radu M. *Dezbateri pe marginea problemei corp-minte*. București, Academia de Științe Sociale și Politice, Academy of Social and Political Sciences, 1984.** [Aceasta este o introducere documentară în problema corp-minte. În ciuda motivelor oarecum conjuncturale care m-au determinat să o apăr, țin în continuare la distincția între problema corp-minte și problema raportului materie-spirit (plasată atunci, în mod oficial, în postura de "problemă fundamentală" a filosofiei). Cred în continuare că sunt posibile combinații derutante între soluțiile la aceste probleme. Lucrarea se concentra în mare parte asupra punctului de vedere Popper-Eccles. A doua parte prezenta fizicalismul. A treia introduce punctele de vedere ale unor autori atât de influenți precum Ryle, Smart sau Quine. Finalul explorează impactul inteligenței artificiale și cercetările referitoare la capacitatea gorilelor de a învăța limbajul surdo-mușilor americani.]
- **STRAWSON, Galen. *Mental Reality*. The MIT Press, 1994.** [Critică neobehaviorismului. Cultivă un *cartezianism naturalizat*, care combină respectul pentru materialism cu ideea că trăirile conștiente sunt singurele fenomene mentale specifice. Discută problema corp-minte. Trece apoi la

intenționalitate și la analiza conceptului de comportament. În final susține că din studiul inteligenței artificiale trebuie trase concluzii opuse celor la care se ajunge în mod obișnuit. Inteligența artificială ne arată, după Galen Strawson, că aria mentalului este mai restrânsă decât se credea, din moment ce sisteme care nu au trăiri conștiente au capacități pe care le credeam mentale (cf. p. 318). Cartea trece cu dibăcie de la probleme tradiționale la cele pe care le-am pune aici la rubricile "mașinăria minții", intenționalitate sau reflecție filosofică asupra inteligenței artificiale.]

- **WILLIAMS, Bernard. *Problems of the Self*. Cambridge University Press, 1973.** [Volum celebru probabil pentru analizele *identității personale*.]
- **WITTGENSTEIN, Ludwig. *Philosophische Untersuchungen*. Oxford, Blackwell, 1953. [Originalul german și traducere în engleză.]** [Wittgenstein, scrie Daniel Dennett, este ostil ideii de investigație științifică a minții (lucru valabil și pentru Ryle). Înțelepciunea standard în știința cogniției este că noi am reușit să trecem dincolo de analizele lor neștiințifice ale minții. Lucrul acesta nu este adevărat, adaugă Dennett. Trebuie să le tolerăm ignoranța în materie de știință și să reținem observațiile lor profunde referitoare la minte, pe care mulți reușesc abia astăzi să le înțeleagă. (cf. Daniel Dennett, "Cog as a thought experiment" URL = <http://ase.tufts.edu/cogstud/papers/cogthogt.htm>) Aceleași idei apar și în Dennett "Further Reading", *Kinds of Minds*.]

### Sistemele computaționale și fundamentele lor teoretice

- **BIHAN, Patrice. *Turbo Prolog. An Introduction to Artificial Intelligence*. Wiley, 1987 [traducere din franceză].** [Ar trebui precizat că Turbo Prolog® este creația unei firme daneze. Produsul a fost distribuit de firma Borland, care l-a plasat în aceeași clasă cu Turbo Pascal® sau Turbo C®. Firma care l-a creat a lansat acum Visual Prolog®, care este disponibil pe Internet ([www.pdc.dk/vip](http://www.pdc.dk/vip))

La respectiva adresă de Internet găsiți și un manual (primele trei capitole, în special, conțin exemple intuitive foarte bune).]

- **BINZIGER, Thomas.** *Jetzt Lerne Ich Delphi. Windows-Programmierung mit Object Pascal.* Markt & Technik Buch- und Software-Verlag, 1996. [Există și o traducere în maghiară (Budapesta: Kossuth, 1998). Marele avantaj al acestei cărți este că pornește de la zero. Se poate învăța Pascal cu ea în mediul de programare Delphi. Acest mediu este de tip RAD: obiecte întregi pot fi integrate în program prin simpla deplasare a unui desen pe ecran. Este pentru cei care îndrăgesc spectacolul interfețelor grafice. Dezavantajul este că mediul Delphi evoluează rapid. În 1999 există deja Delphi 5.0. Trebuie vizitat [www.borland.com](http://www.borland.com) pentru a vedea noutățile.]
- **BROOKS, Rodney.** "Intelligence without representation" în *Artificial Intelligence* 47 (1991), pp. 139-159. [Rodney Brooks, alături de Lynn Andrea Stein, este cel care a condus la MIT echipa care a construit robotul Cog, descris și în Dennett, *Kinds of Minds* (pp. 20-21). În articol el explică de ce el crede că ideea sistemelor cu arhitectură de tip centru-periferie, în care subsistemele comunică între ele prin reprezentări, nu duce nicăieri. Sistemul inteligent propus de Brooks este compus din subsisteme care realizează activități în paralel și care sunt în contact direct cu lumea prin percepție și acțiune. Nu există subsisteme centrale și periferice în această arhitectură.]
- **DEAN, Thomas, James Allen și Yiannis Aloimonos.** *Artificial Intelligence.* Addison-Wesley, 1995. [Introducere bazată pe Lisp în inteligența artificială.]
- **DENNETT, Daniel C.** *Darwin's Dangerous Idea.* Simon & Schuster, 1995. [O lectură cognitivă a darwinismului. A se vedea în special "Natural selection as an algorithmic process" (p. 48 ș.u.).<sup>276</sup> Ideea este că evoluția este un

---

<sup>276</sup> Un exemplu de proces algoritmic este și cupa mondială la fotbal în curs de desfășurare în Franța în momentul în care extrag acest citat din Daniel Dennett. Există un input (cele 32 de echipe participante la turneul din Franța) și un output (campiona lunii). Cu toți putem defini fazele eliminării

proces algoritmic și noi suntem produsul acestui proces. fără ca evoluția să fie un algoritm menit în mod expres să ne genereze. *Ideea periculoasă a lui Darwin* e că nivelul algoritmic este nivelul la care putem cel mai bine să explicăm caracteristicile ființelor vii (cf. p. 59). Cu alte cuvinte, provocarea pe care ne-o lansează cartea este aceea de a gândi evoluția tot în termeni computaționali.]

- **GABBAY, D. și F. Guenther (editori). *Handbook of Philosophical Logic. Volume IV: Topics in the Philosophy of Language*. Reidel, 1989.** [În special partea referitoare la atitudini propoziționale. Capitolul respectiv conține și câteva informații cu caracter istoric, cu privire la felul în care sintagma *atitudini propoziționale* a apărut în opera lui Russell, pentru asemenea "lucruri mentale", precum dorințele, crezămintele, speranțele.]
- **GIARRATANO, Joseph și Gary Riley. *Expert Systems*. Boston, PWS-Kent, 1989.** [Conține și o introducere în limbajul Clips. Surcele în C ale acestui limbaj sunt acum disponibile pe Internet. Vezi și aici despre sistemele expert și Clips.]
- **GILBERT, Stephen și Bill McCarthy. *Object-Oriented Programming in Java*. Waite Group Press, 1997.** [Principala caracteristică a acestei cărți este aceea că nu presupune familiarizarea prealabilă cu programarea. Multe cărți despre Java – limbajul Internet-ului – sunt greu accesibile. În plus este o meticuloasă introducere în programarea orientată pe obiecte. Ar trebui adăugat că uneltele necesare pentru a programa în Java sunt incluse pe CD-ul anexat. Pot fi găsite, de altfel, și pe Internet.]
- **GRASSMANN, Winfried Karl și Jean-Paul Tremblay. *Logic and Discrete Mathematics. A Computer Science Perspective*. Prentice-Hall, 1996.** [Ce anume din logică este interesant pentru știința computerelor? Autorii tratează calculul propozițional, calculul cu predicate, funcțiile recursive, mulțimile și relațiile. Un capitol special este dedicat limbajului Prolog. Este tratat și lambda calculul. Restul cuprinde matematica discretă (grafuri, arbori etc.).

---

diverselor echipe, desena arborele întâlnirilor ș.a.m.d., nu putem, desigur, și cu certitudine dinainte rezultatul. Algoritmul ne este însă deosebit de clar.

gramatici, specificarea formală a programelor, demonstrarea corectitudinii programelor, bazele relaționale de date. Manualul presupune cunoașterea limbajului Pascal.]

- **JAMSA, Kris și Lars Klander. Totul despre C și C++. București, Teora, 1999.** [C și C++ sunt limbaje pentru programatori. De C/C++ este nevoie însă și dacă vrei să-ți faci o idee "de la sursă" despre rețele neuronale și multe altele. Sunt zone deci ale științei cogniției în care se poate pătrunde numai cu C++. Problema a fost, când am alcătuit această bibliografie și nu numai atunci, că nu am găsit o carte de C și/sau C++ care să ia lucrurile de la zero, să înainteze pas cu pas și să conțină și suficiente elemente de natură teoretică. Bănuiesc că trebuie să existe. Abia spre sfârșit am găsit această carte. Este foarte amănunțită. În principiu ia lucrurile de la zero, dar cred că o familiarizare cu programarea este tare nimerită. Teoria este expusă intuitiv. Nu conține tipuri abstracte (de date, proceduri, obiecte) precum Eliott B. Kauffman *Pascal* (Reading: Addison-Wesley, 1992), Main și Savitch *Data Structures* sau alte manuale de Pascal, care ar fi o bună introducere prealabilă. Un mare avantaj este includerea pe CD-ul anexat cărții a mediului didactic de programare Borland Turbo C++ Lite, care dispune de tot ce este necesar pentru a crea programe C/C++ cu scopuri educaționale. A se vedea, de asemenea, Kris Jamsa *Succes cu C++* (București: ALL, 1997). Avantajul cărții acesteia a lui Jamsa este că e foarte clară, practică, mai concisă, cu o bună perspectivă asupra obiectelor în sensul programării.]
- **KAMIN, Samuel N. *Programming Languages*. Addison-Wesley, 1990.** [O introducere în universul limbajelor de programare. Mai multă teorie decât în Mak, *Writing Compilers*.]
- **KOLODNER, Janet. *Case-Based Reasoning*. San Mateo, Ca., Morgan Kaufmann Publishers, 1993.** [O introducere în programarea bazată pe cazuri.]
- **KOWALSKI, Robert. "Logic as computer language" în K. L. Clark și S.-A. Tärnlund (editori), *Logic Programming*. Londra, Academic Press, pp. 3-16.**

[Logica își are originea în formalizarea limbii naturale și a raționării umane, susține Kowalski. Utilizarea ei pentru deducția automată a transformat-o și într-un limbaj de programare (cunoscut sub numele de Prolog). Avantajele se văd în construirea de baze de cunoștințe. Baza de cunoștințe este o mulțime de clauze Horn (propoziții atomice sau implicații de forma " $q$  DACĂ  $p_1$  ȘI  $p_2$  ȘI ...  $p_n$ ") Invocarea programului se face prin *interogarea* bazei de cunoștințe. Rezolvarea de probleme este făcută într-un mod apropiat de cel uman (Kowalski ilustrează ideea cu un mic program care rezolvă cunoscuta problemă cu țaranul, lupul, capra și varza). În afara studiului lui Kowalski, volumul cuprinde o introducere în istoria programării logice, aplicații, programe Prolog pentru înțelegerea limbii naturale, probleme tehnice ale dezvoltării logicii ca limbaj de computer, raporturile cu Lisp.]

- **LEAKE, David B. (editor). *Case-Based Reasoning*. AAAI Press/ The MIT Press, 1996.** [Cu contribuții semnate de Roger Schank, Janet Kolodner ș.a.]
- **MAIN, Michael și Walter Savitch. *Data Structures and Other Objects*. The Benjamin/Cummings Publishing Company, 1995.** [Prezintă tehnicile programării orientate pe obiecte în Pascal. Este un curs avansat de știința computerelor, genul de curs de care ai nevoie dacă vrei să știi mai multe despre sistemele computaționale. Dezavantajul este sugerat imediat de anul apariției: Pasionații interfețelor grafice ademenitoare vor fi dezamăgiți.]
- **MAK, Ronald. *Writing Compilers and Interpreters*. Wiley, 1996.** [Dacă inventați un limbaj de programare, aveți în mod sigur nevoie de un astfel de manual.]
- **MEDSKER, Larry R. *Hybrid Intelligent Systems*. Kluwer Academic Publishers, 1995.** [Combinații între sisteme expert, rețele neuronale și așa mai departe.]
- **MIROIU, Adrian. *Introducere în logica filosofică. Logică și formalizare*. Editura Universității din București, 1994.** [Autorul tratează logica filosofică din perspectiva uneltelor (logice) folosite în filosofie. Este evident că materialul din această carte, de la distincția dintre sintaxă și semantică

(p.32), atitudini propoziționale (p.67 ș.u.) până la tehnicile formalizării, ar fi foarte util să fie cunoscut.]

- **POPA, Cornel. *Logica simbolică și bazele de cunoștințe*. București, Universitatea "Politehnica", 1998. [Manualul este legat în două volume separate, dar paginile sunt numerotate în continuare (I: 1-296; II: 297-544).]** [Discută obiectul logicii, legile gândirii. Introduce logica propozițională și logica predicatelor dimpreună cu bogate incursiuni în universul programelor Prolog. Descrie foarte amănunțit metodele de decizie în logică. Se face o conexiune între definiții și bazele de cunoștințe. Conține și reconstrucții în Prolog de spețe (echivalentul în dreptul continental al cazurilor anglo-saxone).]
- **POPA, Cornel. *Teoria acțiunii și logica formală*. Editura științifică și enciclopedică, 1984.** [Teoria formală a acțiunii este relevantă nu doar pentru filosofie. Între altele, ea este importantă și ca o bază teoretică pentru inteligența artificială. Dincolo de aceste considerente, ar trebui amintit faptul că în spatele cărții stau ani mulți de reuniuni la Cercul de logică și praxeologie din Politehnica bucureșteană, condus de Cornel Popa. Chiar dacă nu folosea denumirea de *știința cogniției*, retrospectiv mi se pare limpede că prin cercetările sale Cornel Popa era pionierul acestui tip de investigații la noi. La acest cerc a luat, de altfel, contact și subsemnatul cu ideea legăturilor dintre logică, teoria acțiunii și inteligența artificială. Astăzi cercul și-a înscris în mod explicit știința cogniției pe agenda sa.]
- **PRATT, Terence. *Pascal. A New Introduction to Computer Science*. Prentice Hall, 1990.** [A învăța Pascal sau alt limbaj de programare înseamnă, de fapt, a te familiariza cu un sistem computațional. Acesta există fie ca mașină virtuală, fie ca mașină fizică. Laboratoarele de inteligență artificială folosesc, de pildă, mașini Lisp, pentru a crește viteza proceselor computaționale. Cartea lui Pratt este tipul de lucrare scrisă cu grija de a nu pierde nici un detaliu. El este și autorul lucrării *Programming Languages* (Prentice Hall, 1975), o introducere teoretică însoțită de descrierea unora dintre limbajele principale.]



- **RAO, Valluru B. și Hayagriva V. Rao. C++ Neural Networks and Fuzzy Logic. MIS Press, 1993.** [Rețelele neuronale pot fi simulate pe calculatoare clasice. În felul acesta, un PC poate fi folosit pentru experimente cu rețele.]
- **SCHILDT, Herb. Advanced Turbo Prolog. Osborne McGraw-Hill, 1987.** [Este, de fapt, o introducere în inteligența artificială în care programele sunt scrise în Prolog. Vezi și Bihan *Turbo Prolog* pentru o introducere elementară în Turbo Prolog.]
- **STERLING, Leon și Ehud Shapiro. The Art of Prolog. Ediția a doua. The MIT Press, 1994.** [Cuprinde fundamentele Prolog-ului, o descriere a limbajului Prolog, tehnicile de bază și tehnicile avansate ale programării în Prolog, aplicații, bibliografie. Cartea este cel mai bun punct de plecare în explorarea lumii programării logice. Programele din carte sunt disponibile pe Internet la MIT Press. Pentru a rula programe Prolog este nevoie desigur de interpretoare și/sau compilatoare (vezi Bihan *Turbo Prolog* pentru lucrul cu Turbo Prolog®). Experimentele cu Prolog în lumea academică folosesc, de multe ori, produse necomerciale, cum ar fi SWI-Prolog, sub Unix. Pentru manuale de SWI-Prolog vizitați URL = [www.swi.psy.nl/usr/jan/SWI-Prolog/Manual](http://www.swi.psy.nl/usr/jan/SWI-Prolog/Manual) sau încercați URL = <ftp://swi.psy.uva.nl> (145.18.114.17). O altă soluție este recursul, sub Windows™, la Strawberry Prolog, care este gratis în varianta neprofesională (URL = <http://www.dobrev.com>).]
- **TANNENBAUM, Andrew S. Organizarea structurată a calculatoarelor. Ediția a IV-a. București, Computer Agora Press, 1999.** [Aceasta este, după mine, cea mai bună introducere în arhitectura sistemelor computaționale.]
- **TANNENBAUM, Andrew și Albert S. Woodhall. Operating Systems. Ediția a doua. Prentice-Hall, 1997.** [Aceasta ar fi cartea mea favorită despre sistemele de operare, nu în ultimul rând din pricină că are un caracter practic, pe CD-ul anexat (sau pe Internet) găsindu-se sistemul de operare Minix. Acesta pare a fi sistemul ideal pentru învățarea modului de funcționare a unui sistem de operare. Sistemul este deschis, se poate consulta sursa sa.]

- **TURING, Alan.** "Computing Machinery and Intelligence" în *MIND*, vol.LIX, nr.236 (oct.1950), pp.433-460. [reprodus și în anexă la Robert J. Fogelin, *Understanding Arguments* (New York: Harcourt Brace Jovanovich, 1982), pp. 360-381; traducere prescurtată în românește, cu titlul "Poate oare mașina să gândească?" în *Probleme filozofice și sociale ale ciberneticii* (București: Editura Politică, 1963), pp. 305-330] [Articolul acesta stă, în fond, la originile inteligenței artificiale. Importanța sa crucială nu are nevoie de prea multe sublinieri. Ar fi interesant de observat doar că traducerea românească a apărut într-un volum care conține, cum sugerează și titlul, contribuții din direcția ciberneticii. Astăzi am identifica în cibernetică un precursor al conexiunismului. Chiar dacă este adevărat că și lui Turing i-a venit ideea de a nu introduce explicit programul în mașină, pe Turing l-am asocia mai degrabă cu computaționalismul actual. Pentru diverse aspecte ale operei lui Turing vezi site-ul web <http://www.turing.org.uk/turing/Turing.html> ]
- **WALKER, Henry M.** *The Limits of Computing*. Boston, Jones and Bartlett Publishers, 1994. [Tratează nu doar nedemonstrabilitatea, ci și problemele ridicate de hard și corectitudinea programelor, de complexitate și de către factorii umani. Cartea lui Walker are avantajul de a fi scrisă simplu, clar și colorat, fără o jertfire excesivă a detaliilor de ordin tehnic. De ea sau de ceva asemănător este nevoie pentru a nu pătrunde dezarmat în lumea argumentelor care recurg la problemele insolubile, care au fascinat atât de multe minți. Vezi, în acest sens, o parte dintre criticile la adresa abordării computaționale a minții.]
- **WELSTEAD, Stephen.** *Neural Network and Fuzzy Logic Applications in C/C++*. Wiley, 1994. [Cartea descrie practic modul în care pot fi scrise programe care simulează rețele antrenate prin retropropagare, cum pot fi clasificate formele, cum pot fi modelate haosul și sistemele dinamice nelineare, cum se pot face previziunile financiare, cum funcționează rețelele cu autoorganizare și altele. Cei interesați de logicele vagi, a căror origine se găsește în fond

în interesul filosofilor pentru logici cu mai multe valori, pot găsi, de asemenea, material pentru aplicații practice.]

- **ZEIDENBERG, Matthew. *Neural Networks Models in Artificial Intelligence*. New York, Ellis Horwood, 1990.** [Introducere în sistemele bazate pe rețele neuronale, așa cum arătau ele în faza când au obținut o semnificativă recunoaștere în inteligența artificială.]

### Mașinăria minții

Titlul acestui paragraf este în mod deliberat metaforic. Variatele curente din psihologie, inteligența computațională, lingvistică, neuroștiință sau filosofie au tendințe uneori diametral opuse atunci când este vorba despre *structurile mentale*. Percepția, memoria, limbajul, atenția, gândirea, învățarea, imaginația pot fi nu doar înțelese foarte diferit, ci sunt plasate la niveluri diferite, integrate în moduri contrastante sau chiar pot lipsi cu totul.

În carte, interesul nostru s-a limitat la unele aspecte ale investigării vederii, memoriei și limbajului, susceptibile să ilustreze ideea abordării computaționale a minții. Aici sunt adăugate și titluri care tratează alte sisteme ale minții umane.

- **BODEN, Margaret A. *The Creative Mind: Myths and Mechanisms*. Ediție extinsă. Londra, Abacus, 1991.** [O încercare de abordare computațională a creativității. Distinge două tipuri de creativitate: imposibilistă și improbabilistă. Prima poate fi gândită în termenii abordării clasice computaționaliste. A doua în termenii conexionismului.]
- **BOGDAN, Radu J. *Grounds for Cognition*. Lawrence Erlbaum, 1994.** [Tradusă în limba română la editura ALL. Pornește de la selecția naturală a unor ființe care au scopuri. Orientarea către un scop face ca dezvoltarea cogniției să devină un avantaj. Ființele devin *bune să facă un anume lucru*. Dar aceasta înseamnă că sunt capabile să îndeplinească o anumită *funcție*. Selecția mecanismelor funcționale adecvate devine astfel temeiul cogniției.]
- **COHEN, Gillian, Kiss și Martin Le Voi. *Memory: A Cognitive Approach*. Open University Press, 1993.**

[Prima jumătate a cărții trece în revistă rezultate empirice. Cea de a doua cuprinde modele ale memoriei. Aceste modele folosesc fie structuri de date, ca în programarea computaționalistă, fie rețele neuronale.]

- **CUMMINS, Robert.** "Systematicity" în *The Journal of Philosophy* XCIII, 12/1996, pp. 591-614. [Comentează cunoscuta idee a lui Fodor și Pylyshyn. Cummins crede că trebuie separate chestiunile privitoare la limbaj de cele privitoare la gândire. Cazul celei din urmă ar fi mai problematic.]
- **DRETSKE, Fred.** *Naturalizing the Mind*. The MIT Press, 1995. [Idea *naturalizării* este tot mai răspândită în filosofie. Dretske trece în revistă filosofic mașinăria minții: reprezentări (naturale și convenționale), introspecție, qualia, conștiința, superveniența și externalismul.]
- **FODOR, Jerry și Zenon W. Pylyshyn.** "Connectionism and cognitive architecture: A critical analysis" în *Cognition* 28 (1988), pp. 3-71. [Principalul atac contra conexionismului. Replica principală a venit din partea lui Paul Smolensky. Vezi și aici mai pe larg despre această dispută.]
- **GORDON, Ian I.** *Theories of visual perception*. Ediția a doua. Wiley and Sons, 1996. [Gordon prezintă, pe rând, concepția gestaltistă, abordarea lui Brunswick, abordarea neurofiziologică, perspectiva datelor senzoriale, abordarea ecologică și cea computațională.]
- **GREENE, Judith.** *Language Understanding: A Cognitive Approach*. Open University Press, 1986. [Sunt prezentate sintaxa și semantica în stil computaționalist. Introducere mult mai simplă decât Winograd *Language as Cognitive Process*, dar care cuprinde și semantica.]
- **HORGAN, Terence și John Tienson.** *Connectionism and the Philosophy of Mind*. Kluwer Academic Publishers, 1991. [Pe lângă trecerea în revistă a disputei cu punctul de vedere computaționalist, cartea cuprinde și dezbaterea versiunii conexioniste a mașinăriei minții, cu discuții despre reprezentări structurate sintactic. Punctul de vedere radical

este cel al lui Van Gelder. Nota accentuat critică este conferită de intervenția lui Jerry Fodor și Brian McLaughlin.]

- **HORGAN, Terence.** "Connectionsism and the philosophical foundations of cognitive science" în *Metaphilosophy* 28, 1&2/1997, pp.1-29. [Trecere în revistă a modelelor conexioniste.]
- **JOHNSON-LAIRD, Philip.** "A model theory of induction", *International Studies in the Philosophy of Science*, vol. 8, nr. 1/1994, pp.5-29. [Autorul aplică teoria din cartea sa *Mental Models* (1983) la cazul inducției. Cunoașterea este reprezentată în modele, iar inferențele deductive pot funcționa fără reguli explicite. Concluzia este testată pe model. Inducția este adăugare de informație la un model mental. Restul numărului cuprinde zece comentarii la studiul lui Johnson-Laird și răspunsurile autorului.]
- **KOSSLYN, Stephen M. și Daniel N. Osherson (editori).** *Visual Cognition*. MIT Press, 1995. [Acesta este volumul al doilea din seria *An Invitation to Cognitive Science* editată de Osherson. Cuprinde studii despre reprezentarea suprafețelor, atenția vizuală, recunoașterea vizuală, mișcările ochiului. De asemenea, Kosslyn tratează imagistica mintală. Dretske scrie despre percepția cu sens. Celelalte volume din această serie sunt consacrate limbajului, gândirii și fundamentelor conceptuale ale științei cogniției.]
- **MARR, David.** *Vision*. New York, Freeman and Company, 1982. [Un text simptomatic pentru perspectiva clasică, zisă și computaționalistă. în știința cogniției.]
- **MATTHEI, Edward și Thomas Roeper.** *Understanding and Producing Speech*. Collins, 1983. [Traducere în franceză cu titlul *Introduction à la psycholinguistique* (Paris: Bordas, 1988).] [Structuri gramaticale și structuri mentale. Introducere personală, dar foarte clară, accesibilă.]
- **McCLELLAND, J. L., D. E. Rumelhart și grupul de cercetare PDP.** *Parallel Distributed Processing*, 2 volume. The MIT Press, 1986. [Lucrarea care a relansat conexionismul. Din punct de vedere strict tehnic, au contat soluțiile găsite pentru antrenarea rețelelor. Cuprinde însă și

analize faimoase ale modurilor de a reprezenta, cu accent evident pe superioritatea conexionismului, modele ale subsistemelor organismului, ale vederii etc.]

- **MICLEA, Mircea. *Psihologie cognitivă*. Cluj-Napoca, Gloria, 1994.** [Tratează prelucrarea informației vizuale, atenția vizuală, categorizarea, imagistica mentală, memoria, reprezentarea cunoașterii și decizia.]
- **MINSKY, Marvin L. "Matter, Mind and Models" în Minsky (editor), *Semantic Information Processing*. The MIT Press, 1968.** [URL = <http://www.medg.les.mit.edu/people/doyle/gallery/minsky/mmm.html> (accesat 28.04.98)] [Minsky susține că raporturile dintre fizic și mental cer dezvoltarea unor unelte puternice. Acestea devin disponibile pe măsură ce se lucrează la programe de computer tot mai inteligente. Hofstadter este de părere că acesta e un articol memorabil, scris de unul dintre cei mai profunzi gânditori.]
- **MINSKY, Marvin. *The Society of Mind*. Londra, Heinemann, 1987.** [Fiecare minte, după Minsky, este compusă din procese mai mici. El numește aceste procese agenți. Fiecare agent mental, luat separat, poate face doar ceva simplu, pentru care nu este nevoie de minte (cf. p. 17). Mintea este o combinație de asemenea agenți elementari. Minsky formulează cum nu se poate mai clar ceea ce am putea denumi *antivitalismul* științei cogniției.]
- **POSNER, Michael I. și Marcus E. Raichle. *Images of the Mind*. New York, Scientific American Library, 1994.** [Posner este psiholog și Raichle neurolog. Volumul explorează fascinantă lume a tehnicilor de obținere de imagini ale creierului în funcțiune. Aceste investigații au devenit posibile la mijlocul anilor '1980. Imagini disponibile la URL = <http://www.soton.ac.uk/~coglab/PET/> Vezi și URL = [hebb.uoregon.edu/brainlab/online\\_papers/](http://hebb.uoregon.edu/brainlab/online_papers/)]
- **PYLYSHYN, Zenon. *Computation and Cognition*. The MIT Press, 1986.** [Dacă secretul mașinării minții este cogniția, ce este ea? Iată răspunsul lui Pylyshyn: " ... unul dintre lucrurile principale pe care cei/cele care dispun de cogniție le au în comun este faptul că acționează pe baza reprezentărilor" (p.Xii). Mințile lor au reprezentări ale

mediului înconjurător, ale altor minți ș.a.m.d. Cartea lui Pylyshyn are reputația de a fi o trecere în revistă completă a mașinării minții din punct de vedere computaționalist.]

- **ROLLS, Edmund T. *The Brain and Emotions*. Oxford University Press, 1998.** [Describe natura și funcțiile emoțiilor. Conține o teorie a conștiinței și aplicații la înțelegerea emoțiilor. Folosește rețele neuronale pentru investigarea emoțiilor. După Rolls, emoțiile sunt stări provocate de răsplată și pedeapsă.]
- **Roth, Ilona și John P. Frisby. *Perception and Representation*. The Open University, 1986.** [Partea I cuprinde o introducere în abordarea cognitivă a reprezentării conceptelor. Partea a doua este dedicată percepției. Partea a treia este consacrată teoriei computaționale a vederii.]
- **SCHACTER, Daniel L. și Endel Tulving. *Memory Systems*. MIT Press, 1994.** [Autorii disting cinci mari sisteme ale învățării și memoriei umane: procedural, de reprezentare perceptuală, semantic, primar și episodic.]
- **SIMON, Herbert A. *Models of Thought*. Volumul I, 1979. Volumul II, 1989.** [Orice bibliografie de acest gen ar fi bizară dacă nu ar include trimiteri la scrieri ale lui Herbert A. Simon. Volumele despre modelele gândirii sunt menite să însoțească *Human Problem Solving*. Temele sunt: structurile memoriei, procesele de învățare, rezolvarea de probleme, inducția și formarea de concepte, percepția și înțelegerea. Aceste teme conturează, de altfel, destul de precis ce anume cuprinde mașinaria minții. O parte prețioasă a cărții o constituie descrierile experimentelor cu programe de calculator, de genul programului BACON, destinat generării de teorii din date empirice.]
- **STICH, Stephen. *Deconstructing the Mind*. Oxford University Press, 1996.** [Autorul demontează filosofic mașinaria minții. Cartea ilustrează excelent ideea folosirii noilor unelte oferite de programare pentru a tranșa probleme filosofice. De exemplu, Stich folosește modelele conexioniste de reprezentare distribuită a cunoașterii pentru a arăta că, dacă aceste modele sunt adevărate, atunci recursul la atitudini propoziționale este în dificultate. În

ansamblu, lucrarea tratează psihologia populară și conexiunea, înțelegerea altor minți, intenționalitatea și naturalismul.]

- **THAGARD, Paul. *Computational Philosophy of Science*. The MIT Press, 1988.** [Describe succint, în anexe, structura programelor scrise în Lisp și reprezentarea cunoașterii în acest limbaj. Cartea se bazează pe un program denumit PI (de la "processes of induction"), scris în Lisp. Pentru o altă perspectivă asupra inducției vezi Johnson-Laird, "A model theory of induction".]
- **WINOGRAD, Terry. *Language as a Cognitive Process*. Syntax. Addison-Wesley, 1983.** [O adevărată enciclopedie a abordării clasice computaționaliste a limbajului. Cuprinde descrieri formale ale principalelor tipuri de gramatici ale limbilor naturale.]

### Viziunea computațională asupra minții: dezbaterile critice

- **BAARS, Bernard J. "Can Physics Provide a Theory of Consciousness?",** recenzie la Penrose, *Shadows*, în **PSYCHE**, vol.2, mai 1995. URL = <http://psyche.cs.monash.edu.au/v2/psyche-2-08-baars.html>  
[Baars respinge ideea că problema conștiinței ar putea fi rezolvată cu ajutorul fizicii. După părerea sa, literatura consacrată conștiinței, în ultimul deceniu, ar putea fi baza unei sinteze ce poate fi luată în considerare. Penrose nici nu discută însă această literatură.]
- **CHALMERS, David (managing editor). "Symposium on Roger Penrose's *Shadows of the Mind*",** **PSYCHE. An interdisciplinary journal of research on consciousness**, vol.2. URL = <http://psyche.cs.monash.edu.au/psyche-index-v2.html> [Conține nouă analize ale cărții lui Penrose. Răspunsul lui Penrose este foarte substanțial și se intitulează în mod semnificativ "Beyond the Doubting of a Shadow".]
- **DENNETT, Daniel. Recenzie la *The Emperor's New Mind* in *The Times Literary Supplement*, 29 septembrie – 5 octombrie 1989.** URL = [www.tufts.edu/as/cogstud](http://www.tufts.edu/as/cogstud)



[/papers/penrose.html](#) [Dennett se întreabă care vor fi fost intențiile lui Penrose atunci când a scris cartea. Penrose este în dezacord cu colegii care au înțeles deja mecanica cuantică. El recurge la explicații populare pentru a-și apăra punctele de vedere. Pe cine speră să ralieze în acest caz? Este oare vorba de o "cruciadă a copiilor" în care speră să-i înroleze pe nespecialiști? Dennett ajunge la concluzia că strategia lui Penrose este de a se preface că se adresează publicului larg pentru a-i putea face pe experți să revadă aspecte fundamentale. Rezultatul este foarte rodnic pentru filosofie, chiar dacă filosofii trebuie să rămână spectatori ai dezbaterii.]

- **FODOR, Jerry. *Concepts. Where Cognitive Science Went Wrong*. Oxford, Clarendon Press, 1998.** [Da, da, ați citit bine! Cel care semnează această critică la adresa științei cogniției este nimeni altul decât Jerry Fodor. Este suficient să vă aruncați un ochi pe bibliografia de față pentru a vedea ce rol a avut opera lui Fodor în dezvoltarea componentei filosofice a științei cogniției. Critica lui Fodor este însă din interior. Fodor examinează și respinge pe rând cele mai importante teorii despre concepte, care intervin adesea în știința cogniției. Propune soluția sa, "atomismul informațional".]
- **GUNDERSON, Keith. "The Imitation Game" in *Mind* LXXIII, nr. 290, aprilie 1964, pp. 234-245.** [reprodus în Keith Gunderson, *Mentality and Machines* (Garden City: Anchor Books, 1971) și în anexă la Robert J. Fogelin, *Understanding Arguments* (New York: Harcourt Brace Jovanovich, 1982).] [Formulează obiecții la adresa testului lui Turing. Nu doar ieșirile unui sistem sunt relevante, ci și modul în care se ajunge la ele. Sugerează că este posibil ca Turing să fi crezut că în viitor se va schimba modul nostru de a ne raporta la mașini.]
- **LUCAS, J. R. "Minds, machines and Gödel" in *Philosophy* XXXVI, 1961, pp. 112-127.** [Un extras amplu este reprodus în Douglas Hofstadter, *Gödel, Escher, Bach. An Eternal Golden Braid* (New York: Basic Books, 1979), pp. 388-390. Hofstadter critică apoi, în câteva etape, argumentul lui Lucas. Vezi și Lucas, *Freedom of the Will*.]

- **LUCAS, J. R. *The Freedom of the Will*. Oxford, Clarendon Press, 1970.** [Lucas reia în finalul cărții argumentul din articolul publicat în 1961. Esența faimoasei formule nedemonstrabile a lui Gödel este autoreferirea. Când este vorba despre sisteme fizice, le cerem acestora să fie conștiente de sine și să ne spună ce pot face și ce nu pot face (și le prindem că nu se pot descurca în mrejele formulei lui Gödel). Orice descriere a minții umane văzută ca sistem fizic ce poate fi conceput ca un sistem formal logic va fi atacabilă cu argumente tip Gödel.]
- **PENROSE, Roger. *Shadows of the Mind*. Oxford University Press, 1994.** [Tradusă și în limba română de către Dana Jalobeanu.]
- **PENROSE, Roger. *The Emperor's New Mind*. Oxford University Press, 1989.** [Tradusă în românește cu titlul *Mintea noastră ...cea de toate zilele*. Aluzia din titlul original este ironică. Trimiterea este, evident, la "hainele cele noi ale împăratului".]
- **PUTNAM, Hilary. *Representation and Reality*. The MIT Press, 1988.** [Cu toate că titlul ne-ar îndemna să clasăm lucrarea în secțiunea care urmează, locul ei este mai degrabă aici. Putnam discută critic funcționalismul, la a cărei apariție a contribuit de altfel substanțial. Cu alte cuvinte, viziunea computaționalistă asupra minții nu poate explica natura stărilor mentale. Acesta este mesajul principal al cărții. Pe lângă aceasta, ea cuprinde pagini foarte importante despre înțeles, reprezentări, conținut restrâns (conceptul cu ajutorul căruia Fodor a încercat să scape din mrejele experimentului lui Putnam cu Pământurile îngemănate), referent și adevăr. Corolarul final este că *funcționalismul*, dacă este corect, implică *behaviorismul*.]
- **SEARLE, John. *The Rediscovery of the Mind*. MIT Press, 1992.** [Volumul evaluează starea actuală a filosofiei minții, rolul conștiinței și al intenționalității. Searle este însă mai degrabă faimos pentru critica sa la adresa "rațiunii cognitive". Dorința lui Searle este să bată cuiul final în coșciugul ideii că mintea ar fi un program de computer (cf. p. XI). S-ar putea ca tocmai aici să fie eroarea lui Searle.

Mintea nu este un program de computer, ci un *proces*. Procesul acesta are, mai mult ca sigur, fire multiple. Este declanșat de mai multe programe. Iar după cum știe oricine a scris vreodată un program cât de cât mai complex programul îl poți examina sintactic, așa cum analizezi un text. Procesul, în schimb, poate decurge în mod neașteptat și, în plus, procesele interacționează, conducând la rezultate și mai curioase uneori. De aceea, este greu de crezut că Searle a bătut cuilul unde trebuia.]

### Reprezentări mentale

- **BAKER, Lynne Rudder. *Explaining Attitudes*. Cambridge University Press, 1995.** [Dezvoltă o concepție *realist practică* despre atitudini propoziționale. Atitudinile nu sunt entități teoretice. Atitudinea depinde de ceea ce *ar face* individul în diverse situații. Recursul la situații contrafactuale diferențiază realismul practic de behaviorism.]
- **BRENTANO, Franz. *Psychologie vom empirischen Standpunkt*. 3 volume. Hamburg, Meiner, 1973.** [Originalul 1874 – un volum; Există o traducere în engleză, *Psychology from an Empirical Standpoint* (Londra: Routledge, 1973) din care primele două capitole pot fi găsite și pe web (URL = [http://home.mira.net/~gaffcam/ phil/brentano.htm](http://home.mira.net/~gaffcam/phil/brentano.htm) ).] [“Mintea poate să gândească despre lucruri care nu există” Aceasta este numită uneori și *problema lui Brentano*. Poți să crezi că ai văzut o zână din povești, fără ca ea să existe fizic. Crezămintele sunt însă *despre* ceva. Filosofii spun că ele au un *continut*. În sens tehnic, *intenționalitate* înseamnă “a fi despre”, a avea un conținut în sensul arătat aici. Reprezentările au intenționalitate. Brentano este cel care a introdus în filosofia contemporană acest sens al intenționalității. Brentano a predat la Universitatea din Viena, între 1874-1894. Printre elevii lui Brentano au fost Meinong, Husserl și Twardowski.]
- **CLARK, Austen. “Seeing and Summing” în *Cognition and Brain Theory* 1/1984, pp. 1-23.** [Teoria lui Marr demonstrează că vederea umană folosește un sistem de reprezentare numerică care combină ingenios semantică

pictorială cu semantică propozițională, susține Clark. Aceste reprezentări nu sunt deci nici pictoriale, nici propoziționale. Dihotomia imagine-cuvinte nu stă în picioare când este vorba despre reprezentări vizuale.]

- **CUMMINS, Robert. *Representations, Targets, and Attitudes*. The MIT Press, 1996.** [Conținutul, forma și implementarea reprezentărilor în creier țin de cercetarea empirică. Definirea conceptului de reprezentare privește filosofia. Vezi și cartea mai veche a acestui autor: *Meaning and Mental Representations* (Cambridge, Mass.: The MIT Press, 1989). Watson, în *Representational Ideas*, îl clasează pe Cummins printre neurofilosofi.]
- **DENNETT, Daniel. *The Intentional Stance*. The MIT Press, 1987.** [De ce să privim mașinăria minții ca pe un sistem intențional? Aș putea lua în serios predicțiile astrologice auzite dimineața la radio. Dacă ele ar funcționa pentru unii oameni – spune Dennett – atunci acei oameni ar fi sisteme astrologice (cf. p. 16). Altfel îi privim ca agenți raționali, dotați cu crezăminte și dorințe. Strategia intențională are o extraordinară putere predictivă.]
- **JACKENDOFF, Ray. *Languages of the Mind*. The MIT Press, 1992.** [O mare parte a efortului de cercetare din știința cogniției s-a concentrat asupra limbajului și vederii. Nu este de mirare deci că lingviștii sunt și ei printre cei care au contribuit la dezvoltarea domeniului. Unii dintre ei, precum Jackendoff, încearcă să ducă discuția și dincolo de limbajul natural. Cartea conține capitole despre cogniția spațiului, muzică și problema realității.]
- **JACOB, Pierre. *What Minds Can Do. Intentionality in a Non-Intentional World*. Cambridge University Press, 1997.** [Prima parte este dedicată naturalizării intenționalității. Cea de a doua tratează problema rolului cauzal al intenționalității. Discută teoria computațională a minții bazată pe reprezentări.]
- **MCGINN, Colin. *Mental Content*. Oxford, Blackwell, 1989.** [Care este locul conținuturilor mentale: "în cap" sau nu? Primul răspuns este cel *internalist*. McGinn discută pe larg varianta sa de externalism.]

- **SEARLE, John.** *Intentionality*. Cambridge University Press, 1983. [Searle însuși socotește că preocuparea pentru intenționalitate este centrală în filosofia sa. Modelul actelor de vorbire este un mod de a vedea intenționalitatea ca reprezentare.]
- **WATSON, Richard.** *Representational Ideas: From Plato to Patricia Churchland*. Kluwer Academic Publishers, 1995. [După cum spune autorul, acesta este un studiu *structural* al teoriilor despre reprezentări în filosofia occidentală. Văzute "de la distanță", aceste teorii ar fi marcate de firul roșu al ideii de *izomorfism*.]

### Nativism, modularitate și limbaj al gândirii

Nu văd de ce ar exista o conexiune *necesară* între aceste teme. În practică însă, Fodor, cel care a lansat și ipoteza modularității și pe cea a limbajului gândirii, adoptă și poziția nativistă, conform căreia, de exemplu, capacitatea de a stăpâni o limbă naturală este *innăscută*, este o *capacitate nativă*, nu una dobândită prin învățare. Karmiloff-Smith, care dezvoltă ideea depășirii modularității, discută și ea implicațiile nativismului. Oricum, nativismul și modularitatea sunt teme care dau specificul filosofiei minții și științei cognitive actuale. De aici și ideea de a separa acest nod tematic de rest.

În altă ordine de idei, nativismul este un excelent punct de plecare pentru înțelegerea cognitivismului (v. celebra recenzie a lui Chomsky la cartea lui Skinner *Verbal Behavior* mai jos).

- **CHOMSKY, Noam.** "A review of B. F. Skinner's *Verbal Behavior*" în *Language* 35, 1/1959, pp. 26-58. [Retipărit în Fodor și Katz *The Structure of Language*, pp. 547-578.] [După Skinner, spune Chomsky, stimulările externe actuale și istoria condiționărilor pozitive sunt suficiente pentru a determina comportamentul verbal. Intervenția vorbitorului ca atare ar fi minimă. Chomsky examinează pe rând conceptele folosite de Skinner cu scopul de a arăta că, luate literal, ele nu acoperă aproape nici un aspect al comportamentului verbal și, luate în sens metaforic, pierd orice avantaj față de formulările tradiționale

După ce ajunge la această concluzie, se referă la studiul lui Lashley despre ordinea serială în comportamente (un moment crucial în trecerea de la behaviorism la cognitivism, la mijlocul secolului al XX-lea). Organizarea sintactică a rostriilor nu transpare în structura fizică a acestora. A construi un enunț, cu alte cuvinte, nu este o operație echivalentă cu înșiruirea unor mărgelile pe un fir. Gramatica specifică enunțurile tot așa cum o teorie deductivă specifică teoremele. Gramatica este complexă. Aici Chomsky face pasul decisiv și arată, urmându-l pe Lashley, că gramatica face parte din comportament. Copilului i-ar fi imposibil să o construiască pe baza stimulărilor externe. Aceste stimulări sunt prea sărace și gramatica prea complexă pentru ca operația să reușească în timp scurt. Concepția către care conduce recenzia este limpede una *mentalistă și nativistă*.]

- **CHOMSKY, Noam. *Cunoașterea limbii*. București, Editura științifică, 1996. [Postfață de Alexandra Cornilescu. Ediția originală: Praeger, 1985.]** [Discută, între altele, ceea ce numește Chomsky "problema lui Platon", problema sărăciei datelor de intrare în sistemul minții și contrastul cu ieșirile foarte bogate și articulate. Postfața ne introduce în concepția lui Chomsky despre limbă, inclusiv în atmosfera ipotezei că lingvistica este o parte a psihologiei cognitive.]
- **CHOMSKY, Noam. *Language and Mind*. Harcourt Brace Jovanovich, 1968.** [Contribuții lingvistice la studiul minții.]
- **FODOR, Jerry A. și Jerrold J. Katz (editori). *The Structure of Language. Readings in the Philosophy of Language*. Prentice-Hall, 1964.** [Am inclus aici această antologie mai veche, deoarece ni se pare că ea este astăzi relevantă mai ales ca proiect de a produce o *filosofie a lingvisticii*. Fodor și Katz resping atât filosofia pozitivistă a limbajului, pe care o acuză de lipsă de control empiric, cât și filosofia limbajului obișnuit, pe care o critică pentru apel la judecăți intuitive și caracter nesistematic (p. 18). Filosofia limbajului urmează să fie o analiză a conceptelor, teoriilor și metodelor lingvisticii.]
- **FODOR, Jerry. "Why there still has to be a language of thought", anexă la Jerry Fodor, *Psychosemantics*. The**

- MIT Press, 1987, pp. 135-154.** [Apără ipoteza limbajului gândirii cu ajutorul conceptelor de structură sintactică, productivitate, sistematicitate.]
- **FODOR, Jerry. *The Language of Thought*. Harvard University Press, 1975.** [Sursa dezbaterii despre limbajul gândirii.]
  - **FODOR, Jerry. *The Modularity of Mind*. The MIT Press, 1983.** [Cu toate că Fodor are o serie de referiri cu caracter istoric, se poate spune că aceasta este cartea care a lansat ipoteza modularității minții.]
  - **KARMILOFF-SMITH, Annette. *Beyond Modularity*. MIT Press, 1992.** [Autoarea a fost studenta și apoi asistenta lui Piaget la Geneva. Cartea pomește de la o critică a modularității în sensul lui Fodor. Dezvoltă însă o teorie care încearcă să reconcilieze nativismul cu constructivismul lui Piaget. Trece în revistă dezvoltarea minții copilului.]
  - **RANTALA, Veikko și Tere Vadén. "Minds as connoting systems: logic and the language of thought" în *Erkenntnis* 46, 3/1997.** [Autorii critică modelul lui Fodor și Pylyshyn. Cred că modelul semiotic propus de Fetzer este mai realist.]

### Interpretarea minților

- **BOGDAN, Radu J. *Interpreting Minds*. MIT Press, 1997.** [Vezi aici capitolul "Noi orizonturi" pentru o prezentare a cărții.]
- **CARRUTHERS, Peter și Peter K. Smith (editori). *Theories of Theories of Mind*. Cambridge University Press, 1996.** [Cartea ilustrează cum nu se poate mai bine regăsirea filosofiei cu psihologia sub cupola științei cognitive. Diviziunile lucrării sunt dedicate, pe rând, contrastului dintre interpretare ca teorie și ca simulare, problemelor legate de modularitate, autism și evoluție.]
- **DAVIES, Martin și Tony Stone. "Folk psychology and mental simulation". URL = [www.nyu.edu/gsas/dept/philo/courses/concepts/folkpsychology.html](http://www.nyu.edu/gsas/dept/philo/courses/concepts/folkpsychology.html) [accesat 28.04.98]** [Urma să apară în *Contemporary Issues in the*

*Philosophy of Mind* editată de Anthony O'Hear la Cambridge University Press. Articolul dezvoltă epistemologia simulării. Discută predicția și explicația cu ajutorul simulării. Literatura relevantă pentru acest studiu, spun autorii, este cuprinsă în linii mari în trei volume: *Folk Psychology: The Theory of Mind Debate* editat de M. Davies și T. Stone (Oxford: Blackwell, 1995), *Mental Simulation: Evaluations and Applications* editat de M. Davies și T. Stone (Oxford: Blackwell, 1995) și Carruthers și Smith, *Theories of Theories of Mind*.]

- **DAVIES, Martin.** "The Mental Simulation Debate", *Proceedings of the British Academy* vol. 83, 1994, pp. 99-127. [Introducere în istoria disputei privitoare la natura psihologiei naive. Este aceasta o teorie despre procesele din mintea celorlalți? Sau este vorba despre simularea proceselor din mintea celorlalți? Se concentrează apoi asupra simulării.]
- **HAYES, C. M.** "Theory of mind in nonhuman primates" în *Behavioral and Brain Sciences* 21, 1/1998, pp. 101-134. [Sunt puține motive pentru a crede că maimuțele au o teorie a minții celorlalți. Trece în revistă cercetările, începând cu opera de pionierat a lui David Premack și G. Woodruff. Propune o strategie pentru a decide dacă semnele comportamentale care ar putea indica prezența unei teorii a minții altora se datorează unei activități mentale sau nonmentale.]
- **PEACOCKE, Christopher (ed.).** *Objectivity, Simulation and the Unity of Consciousness. Current Issues in the Philosophy of Mind.* Oxford University Press, 1994. [Include și comentariile lui Jane Heal și Joseph Perner la Davies, "Mental Simulation".]
- **PERNER, Joseph.** *Understanding the Representational Mind.* The MIT Press, 1991. [Perner este psiholog. Și-a câștigat un deosebit renume în urma rezultatelor obținute cu testul crezămintelor false. Cartea aceasta este o ilustrare paradigmatică a reîntâlnirii psihologiei cu filosofia. Partea I discută conceptul de *reprezentare*. Partea a doua este dedicată minții în interacțiune cu alte minți: interacțiunile sociale timpurii și primii pași ai unei teorii a minții;



dobândirea unei teorii a cunoașterii; înțelegerea gândirii și a crezămintelor; dorințele. Partea a treia discută probleme ale dezvoltării minții.]

### Conștiința

- **BAARS, Bernard J.** "In the theater of consciousness" în *Journal of Consciousness Studies* 4, 4/1997, pp. 292-309 [URL = <http://www.imprint.co.uk/theatre.htm> ]. [Dezvoltă punctul de vedere al spațiului de lucru global. Vezi și cartea lui Baars *In the Theater of Consciousness: The Workspace of the Mind*. (Oxford University Press, 1997).]
- **BAARS, Bernard J.** *A Cognitive Theory of Consciousness*. Cambridge University Press, 1988. [URL = <http://www.wrightinst.edu/faculty/~baars/book/> ] [Dezvoltă o teorie a conștiinței ca spațiu global de lucru. Baza în creier a acestui spațiu al conștiinței ar fi un Sistem de Activare Extins Reticularo-Talamio (în englezește ERTAS). Sistemele inconștiente sunt contextuale și contrastează cu sistemul conștiinței, care este global. Discută atenția, eul, autoobservarea. Contrar altor autori, Baars crede că sistemul conștiinței are o multitudine de funcții vitale în sistemul nervos.]
- **BLOCK, Ned.** "On a confusion about a function of consciousness" în *Behavioral and Brain Sciences* 18, 2/1995, pp. 227-287. [Varianta preliminară pe web la URL = <http://www.cogsci.soton.ac.uk/bbs/Archive/bbs.block.html> ] [Susține că, de fapt, conceptul de conștiință nu este decât o cupolă sub care sunt puse multe alte "conștiințe".]
- **CARRUTHERS, Peter.** *Language, Thought and Consciousness*. Cambridge University Press, 1996. [Carruthers se opune teoriei comunicaționale a limbajului natural (o ilustrare a acestei teorii ar constitui-o poziția lui Jerry Fodor în *Modularity of Mind*, unde limbajul face parte dintre modulele care asigură intrările și ieșirile din sistem). Investighează ipoteza rolului limbajului în gândire. Propune o teorie a conștiinței bazată pe gândirea reflexivă.]

- **CHALMERS, David.** *On the Search of a Neural Correlate of Consciousness.* URL = <http://www.arizona.u.edu/~chalmers/papers/ncc.html> [Ideile din acest articol sunt reluate în Chalmers, *The Conscious Mind*, cap.6.]
- **CHALMERS, David.** *The Conscious Mind.* Oxford University Press, 1996. [Vezi capitolul anterior pentru o prezentare a cărții.]
- **CONSTANDACHE, G. G. (ed.).** *Cum ne țesem eu.* București, ALL, 1998. [Culegerea este consacrată problemei conștiinței, tratată din punctul de vedere al neuroștiinței, sistemelor dinamice și filosofiei. Eseul lui Dennett, piesa de rezistență a volumului, este cel care dă și titlul antologiei.]
- **CRICK, Francis și Christof Koch.** "Consciousness and neuroscience" în *Cerebral Cortex* 8, 1998, pp. 97-107. [Crick și Koch sunt partizanii existenței corelatului în creier al conștiinței (în engleză NCC). Discută ipoteza rolului oscilațiilor de 30-70 Hz. Resping ideea filosofică a unor ființe zombi. Admit însă că o parte a creierului funcționează zombi. Pe filosofi, spun ei, îi fascinează *qualia*: calitatea de albastru (în engleză *blueness*) a albastrului etc. Se poate explica însă de ce trăirea culorii este privată. În fiecare etapă a procesului vederii, în cortexul vizual, datele explicite sunt recodate. Noi putem distinge între culori în etapele ulterioare, dar nu mai putem relata verbal trăirea.]
- **DENNETT, Daniel.** *Consciousness Explained.* Boston, Little Brown, 1991. [Dezvoltă modelul versiunilor (*drafts*) multiple ca explicație a conștiinței. Dennett respinge recursul la *qualia*.]
- **DENNETT, Daniel.** *Content and Consciousness.* Londra, Routledge and Kegan Paul, 1969. [Aceasta este prima carte a lui Dennett. Începe cu o teorie despre *intenționalitate*. Pe temeiul ei este construită teoria conștiinței. Dennett accentuează asupra centralității limbajului pentru conștiință. Vederile lui Dennett au suferit schimbări în lucrările sale ulterioare.]
- **GALLUP, Gordon.** *Can animals empathize? Yes.* [Apărut inițial în *Scientific American*; disponibil la URL =

<http://www.scientificamerican.com/1998/1198intelligence/1198gallup.html> ] [Pot animalele să recurgă la empatie? Gallup spune că da. Face incursiuni și în lumea pacienților care nu mai pot vedea conștient (*blindsight*). Numeroase alte legături pe web. Vezi și Povinelli *Can animals empathize. Maybe not.*]

- **HAMEROFF, Stuart și Roger Penrose. *Orchestrated Objective Reduction of Quantum Coherence in Brain Microtubes: the "Orch OR" Model for Consciousness*. URL = <http://www.arizona.u.edu/~hameroff/or.html>** [O teorie cuantică a conștiinței. Dezvoltă idei din Penrose, *Shadows*.]
- **JACKSON, Frank. "Epiphenomenal Qualia" în *Philosophical Quarterly*, 32 (1982), pp. 127-36. [URL = <http://members.aol.com/NeoNoetics/Mary.html> ]** [Unul dintre cele mai cunoscute experimente imaginare din filosofia minții: cel cu Mary, cea care trăiește în camera în care totul se vede în alb și negru.]
- **JOHNSON-LAIRD, Philip. "A computational analysis of consciousness" în *Cognition and Brain Theory* 4/1983, pp. 499-508.** [Cheia problemei conștiinței stă în răspunsul la întrebarea ce înseamnă pentru un program să aibă un model de nivel înalt al propriilor operații.]
- **McGINN, Colin. *The Problem of Consciousness*. Oxford, Blackwell, 1991.** [A sosit momentul ca să admitem că nu putem dezlega misterul conștiinței (cf. p. 1). Ca să putem proiecta o mașină conștientă ar trebui să explicăm modul în care este creierul conștient, iar acest lucru nu-l știm (cf. p. 213).]
- **METZINGER, Thomas. *Subjekt und Selbstmodell*. Ediția a doua. Paderborn, Mentis Verlag, 1999.** [În centrul atenției, în carte, de află problema eului. Metzinger, împreună cu Georg Meggle, este editorul, la Mentis Verlag, al seriei **Geist, Erkenntnis, Kommunikation** (URL = [http://www.mentis.de/forum\\_set.html](http://www.mentis.de/forum_set.html) ).]
- **NAGEL, Thomas. "What is it like to be a bat" în *The Philosophical Review* 1974, pp. 435-450 [URL =**

[http://members.aol.com/NeoNoetics/Nagel\\_Bat.html](http://members.aol.com/NeoNoetics/Nagel_Bat.html) ].  
[Care sunt trăirile unui liliac?]

- **NOË, Alva, Luiz Pessoa și Evan Thompson.** “Beyond the grand illusion: what change blindness really teaches us about vision” [urmează să apară în *Visual Cognition* 7, 2000; versiune penultimă la URL = <http://www2.ucsc.edu/people/anoel/GrandIllusion.html> ] [Numai o parte din mediu este codificată în detaliu în condiții normale. Oamenii sunt orbi la aspecte care se schimbă într-o imagine (prin clipiri) dacă acestea nu au fost codificate. Teoria lui Marr, care presupune o reprezentare elaborată a lumii vizibile, nu este plauzibilă, din acest punct de vedere. Nu există, după autori, nici dovezi în favoarea tezei mult mai tari că scenele pe care le vedem sunt o mare iluzie. Trebuie luată în considerare și integrarea vederii cu alte sisteme de explorare a mediului.]
- **NORDBY, Knut.** *Vision in a complete achromat* [URL = <http://www.u.arizona.edu/~chalmers/misc/achromat.html> ] [Autorul are el însuși o vedere acromatică. Relatează experiența sa de viață. Informațiile pe această temă sunt relevante pentru analiza experimentului imaginat de către Frank Jackson.]
- **PALMER, Stephen E.** “Color, consciousness and the isomorphism constraint” în *Behavioral and Brain Sciences* 22, 6/1999 [versiune penultimă pe web la adresa indicată în introducerea la acest capitol]. [Discută psihologia percepției culorilor. Analizează posibilitatea “inversării spectrului” în sensul lui John Locke.]
- **PAUEN, Michael.** *Das Rätsel des Bewußtseins.* Paderborn, Mentis Verlag, 1999. [Extras la URL = <http://staff-www.uni-marburg.de/~pauen/Bewusstsein.htm> ] [Problema aflată în centrul atenției este natura conștiinței fenomenale. Chestiunea cheie ar fi raportul dintre perspectiva persoanei întâi și perspectiva obiectivă a neuroștiinței (o problemă deci de ordin epistemic). Soluția spre care înclină Pauen este o formă a teoriei identității. Prima parte a cărții analizează modul de a pune aceste probleme și avantajul soluțiilor. În partea a doua sunt discutate două perspective: identitatea și

dualismul proprietăților (cu referiri la autori precum Kripke, dar și cu exemple – durere, frică). În altă ordine de idei, trebuie adăugat că, pentru anul 2000, Pauen anunță apariția, unui volum editat împreună cu Gerhard Roth și intitulat *Neuro- und Kognitionswissenschaften* (Paderborn/München: Fink/UTB) și a unei introduceri personale în filosofia minții, *Grundprobleme der Philosophie des Geistes* (Frankfurt: Fischer). Pentru alte publicații de filosofia minții și știința cognției, în germană, la Mentis Verlag din Paderborn, vezi URL = <http://www.mentis.de/> ]

**POVINELLI, Daniel J. *Can animals empathize? Maybe not.* [Apărut inițial în *Scientific American*; disponibil la URL = <http://www.scientificamerican.com/1998/1198intelligence/1198povinelli.html> ]** [Povinelli crede că animalele au mari limite când este vorba despre înțelegerea minții celorlalți. Polemizează cu Gallup]